UNIVERSITY OF LONDON THESIS

Degree PhD Year 2006 Name of Author VINAYAGAMOORTHY, Vinoba

Participant Responses to Virtual Agents in Immersive Virtual Environments

Vinoba Vinayagamoorthy

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of the
University of London.

Department of Computer Science
University College London

July 18, 2006

To my parents

I, Vinoba Vinayagamoorthy, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis is concerned with interaction between people and virtual humans in the context of highly immersive virtual environments (VEs). Empirical studies have shown that virtual humans (agents) with even minimal behavioural capabilities can have a significant emotional impact on participants of immersive virtual environments (IVEs) to the extent that these have been used in studies of mental health issues such as social phobia and paranoia. This thesis focuses on understanding the impact on the responses of people to the behaviour of virtual humans rather than their visual appearance. There are three main research questions addressed.

First, the thesis considers what are the key nonverbal behavioural cues used to portray a specific psychological state. Second, research determines the extent to which the underlying state of a virtual human is recognisable through the display of a key set of cues inferred from the behaviour of real humans. Finally, the degree to which a perceived psychological state in a virtual human invokes responses from participants in immersive virtual environments that are similar to those observed in the physical world is considered.

These research questions were investigated through four experiments. The first experiment focused on the impact of visual fidelity and behavioural complexity on participant responses by implementing a model of gaze behaviour in virtual humans. The results of the study concluded that participants expected more life-like behaviours from more visually realistic virtual humans. The second experiment investigated the detrimental effects on participant responses when interacting with virtual humans with low behavioural complexity. The third experiment investigated the differences in responses of participants to virtual humans perceived to be in varying emotional states. The emotional states of the virtual humans were portrayed using postural and facial cues. Results indicated that posture does play an important role in the portrayal of affect however the behavioural model used in the study did not fully cover the qualities of body movement associated with the emotions studied. The final experiment focused on the portrayal of affect through the quality of body movement such as the speed of gestures.

The effectiveness of the virtual humans was gauged through exploring a variety of participant responses including subjective responses, objective physiological and behavioural measures. The results show that participants are affected and respond to virtual humans in a significant manner provided that an appropriate behavioural model is used.

Acknowledgements

My research was supported by the EPSRC funded IRC project: EQUATOR (Grant GR/N15986/01). This thesis would not have been possible without the help of so many people and it certainly would not have been as enjoyable without them. I wish it were possible to thank them all.

Never in a million years did I imagine the possibility of working in Virtual Reality before meeting my supervisors, Mel Slater and Anthony Steed. I would like to thank Mel Slater, my primary supervisor, for giving me the opportunity to undertake a PhD in such an exciting field. His excellent guidance and support, often beyond the call of duty, has been invaluable in developing my ideas and research skills. I got lucky, a second time, in my secondary supervisor, Anthony Steed. His technical expertise, feedback and attention to detail were instrumental in turning my ideas into practical experiments.

A number of people gave me their expertise and time. Three special colleagues (and good friends) have each played a significant role in helping me realise my research: Andrea Brogni, Marco Gillies and David Swapp. Andrea tirelessly helped me to run most of my experiments, understand physiological devices and appreciate Italian lunch hours. Marco's extensive knowledge and technical skills in virtual human animation was used extensively, and often mercilessly, when I was trying to make my experiments feasible. This research would have definitely taken more effort and energy to complete without his unwavering help. It is my intention to reduce the demands I put upon him from this point on. I would have had to carry out my experiments on a desktop without David's constant advice and technical support on Virtual Reality systems. That would have been a calamity.

I have often persuaded many colleagues (and good friends) of mine to help me run experiments and model virtual environments. Angus Antley deserves a medal for helping me out every time I asked. I have also had much fun dressing up for his Halloween parties. Ashwin Beeharee helped me run experiments even though it had nothing to do with his research. Joel Jordan designed a virtual street for the preliminary experiment on visual realism and introduced me to the Falafel. I would like to take this opportunity to thank Mark Coulson, from Middlesex University, for generously allowing me to use his data, questionnaire and expertise on affective postures. Richard Milton gave his voice to an agent and Mette Thomsen modified and gave me use of her virtual hotel room for the experiment on postures. Mette also taught me how to make a virtual curtain in less than five minutes! I would also like to thank the talented, William Gregory, for convincingly playing the part of a virtual human in the experiment on kinesics. William also helped in teaching my agents how to express Anger and Sadness over closing a bar. Jean-Daniel Nahmias expertly modelled virtual masks for use in the same experiment.

I have been most fortunate to have carried out my PhD with an amazing set of people within the Virtual Environments and Computer Graphics research theme at UCL. I have often used them as sounding boards and can not thank them enough for being good friends throughout this experience. People I have not already mentioned include my fellow PhD students Pankaj Khanna, Jesper Mortensen, and Maria Roussou. I am sure I have missed out many others who have helped me along the way. I thank you too.

In my life away from UCL, Vidula Vinayagamoorthy (Vidu), my sister, flatmate and friend, kept me in good cheer and has often bailed me out of doing tedious work in my PhD. This included transcribing interviews for next to nothing and making copious amounts of caffeinated goodness. This was often followed by lunches and dinners made by my fiancé Rasarethenam Santhirakumar (Chandru). I would like to thank them both for putting up with me through the stressful times and keeping me sane during the periods in which I truly believed my world would end due to a blue screen on my laptop.

Finally, and most importantly, I would like to thank my Mum and Dad for being extremely supportive in so many ways throughout my life. My Mum's belief in me has never fallen and my Dad's generosity has known no bounds. Not only did they bring me into the world, they made sure I got the best of everything in it. I could not have even dreamt of getting this far without their love and encouragement. I will forever be indebted to them. This thesis is dedicated to them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the novel *Snow Crash*, Neil Stephenson envisioned a Virtual Reality (VR) version of the Internet - the Metaverse (Stephenson, 1992). In the Metaverse, Stephenson's virtual humans were as diverse as the people they represented in both appearance and behaviour. Their appearance varied with ethnicity, status and resemblance to the individual and they portrayed personalities, emotional expressiveness, status and life-imitating behaviours. In fact, the virtual humans were so highly advanced that participants often held social gatherings and business conferences in the Metaverse. The beginnings of the Metaverse are observable in online chat worlds, virtual gaming, through to virtual newsreaders. However, this is still a long way from the Metaverse and research presented in this thesis is a contribution to its eventual realisation.

Virtual humans are an essential part of the content in many types of application such as in entertainment, games and story-telling (Arts, 2006; Cavazza et al., 2002; Machado et al., 2001), training environments (Gratch and Marsella, 2005; Johnsen et al., 2005), virtual therapy (Pertaub et al., 2002; Herbelin et al., 2004), conversational representatives (avatars) (Vilhjálmsson and Cassell, 1998; Cassell et al., 1999), and expressive conversational interactive agents (Gillies and Ballin, 2003; Rosis et al., 2003). They are used within human-computer interaction in order to harness the users' automatic responses to the human form and behaviour, and thereby allow users to achieve a kind of empathic interaction that would otherwise be difficult. The ultimate goal is to be able to construct virtual humans to which participants respond as if they were real, despite knowing that they are not. Unfortunately, this is difficult to achieve in practice. There are a variety of tools to aid the avid animator/modeller in the creation of non-interactive animated virtual humans. However, apart from needing artistic talent and a thorough understanding of the human anatomy, the process is time-consuming, expensive and labour-intensive. The creation of an interactive and responsive virtual human for purposes of facilitating activities like communication, learning, training, and therapy in a virtual environment (VE) is a completely different ball game. In the fields of computer graphics, one of the main challenges has been to create a virtual human, that is not only visually lifelike but also capable of replicating real-life context-appropriate behaviours with such high fidelity that it is indistinguishable from one that is fully user-controlled.

One of the problems in creating such a virtual human is that the behavioural cues used in the physical world are so varied and rich in detail that it is often difficult to reduce complex behaviour patterns into an algorithm without in-depth studies. In addition, the advent of graphics cards with more memory, faster texture rendering and hardware acceleration has meant that the process of making a visually enhanced virtual human, is far more advanced than the creation of plausible behaviours. This introduces a tension between the visual appearance of the virtual human and behavioural capabilities it can offer. Evaluative studies have suggested that in scenarios requiring social interaction, feedback and communication within VEs, the more visually realistic the representation gets, the more naturalistic individuals expect the virtual human to act (Tromp et al., 1998; Slater and Steed, 2002). On the other hand, studies have also shown that virtual humans with minimal behavioural and visual fidelity are capable of eliciting significant participant responses (Garau et al., 2005; Pertaub et al., 2001; Slater and Steed, 2002). One approach to designing effective virtual humans is to compare and contrast the responses evoked in participants of VEs with those observed in the physical world. If an individual responds to the virtual human as if it were real despite the knowledge that it is not, then the virtual human is effective (Sanchez-Vives and Slater, 2005).

The research presented in this thesis concerns the design and implementation of affective behavioural cues in virtual humans by mimicking those displayed by individuals in the physical world. In other words, this thesis focuses on building affective virtual humans. The effect of incorporating these behaviours into virtual humans has been studied by conducting controlled experiments designed to observe social interactions between participants and a virtual human in an immersive virtual environment (IVE). The success of the virtual human is measured in terms of the subjective, physiological and behavioural responses of the participant during the interaction.

1.1 Research Challenges

In many of the applications, mentioned above, the demands on a virtual human include the facilitation of effective communication and interaction between the participants since the VE becomes a place of social interaction (Slater and Steed, 2002). This often highlights flaws in the behavioural abilities of the virtual human. In the physical world, individuals interacting with each other draw impressions about the mood and emotional states of their conversational partners (Argyle, 1969). Under similar circumstances in the virtual world, where participants are represented by virtual humans, a lack of expressiveness greatly deteriorates the experience of social interactions (Schroeder, 2002). Gratch et al. (2002) report on the difficulties involved in integrating existing technologies to create a responsive virtual human similar to the ones envisaged by Stephenson (1992) in the Metaverse that is versed in the rules of communication, expression and movement.

On of the main challenges encountered is the lack of emotional expression in virtual humans. This leads participants to judge the virtual human as "*cold*" (Tromp et al., 1998; Slater et al., 2000). In order for virtual humans to invoke realistic responses in participants, it needs to display some level of emotional expression. The creation of emotional expression in virtual humans requires models which *generate* an emotional state and *express* the emotional state using behavioural cues.

There are many models which dynamically generate an underlying emotional state. However, research into creating plausible behavioural cues have been heavily biased towards facial expression since the face is the most observed behavioural cue in social interactions. There is relatively little research conducted on affective kinesics (postural cues and body movement). The reason for this is two-fold. Firstly, the relationship between the body and its role in the communication of affect is less understood in comparison to facial expression. Some researchers believe that the body simply expresses the intensity of an individual's underlying emotional state (Ekman and Friesen, 1967), while others believe that the body can be a more dominant cue to the individual's underlying emotional state (de Gelder, 2006). Secondly, the body has many degrees of freedom which makes it difficult to study. Therefore, there is much less research in social psychology regarding the role of the body in the communication of affect. This presents difficulties in trying build a model to represent affective bodily cues.

Yet, individuals in the physical world use a variety of congruent behavioural cues, including their bodies, to express their emotional states. In interactive applications involving full-body virtual humans, it may be vital to express emotional states through facial expressions and congruent kinesic cues.

1.2 Hypothesis

The main hypothesis of this research is that virtual humans controlled by a model of affective behaviours (kinesic) can elicit appropriate responses in participants of immersive virtual environments. This is mainly associated with the premise that kinesics play a significant role in the communication of affect. The research questions associated with this hypothesis were investigate through four experiments in an IVE.

The idea of using models of behaviours to invoke a response in participants of VEs is not new. This thesis extends existing work and focuses on designing models of behavioural cues which can be used to express a virtual human's underlying emotional state. Even though participants are aware that the virtual human is an artificial creation, previous studies suggest that they will respond to virtual humans as if they were real (Sanchez-Vives and Slater, 2005). The underlying premise is that if a virtual human displays behavioural cues typical of a certain psychological state, participants interacting with the virtual human will respond appropriately. The aim of this research is to portray an emotional state in virtual humans through behavioural cues and investigate if the virtual human elicits realistic participant responses in an interaction.

The two behavioural cues investigated in this thesis are: kinesics (postural attributes and quality of body movement) and gaze behaviour. The communication of affect through kinesics was the main focus of the thesis, however, gaze behaviour was chosen as a first step to investigate the effectiveness of behavioural models.

Two preliminary experiments were designed to investigate the importance of behavioural fidelity in virtual humans. A further two experiments were designed to explore the importance of kinesics in the communication of two affective states: *Anger* and *Sadness*.

1.2.1 Preliminary experiments

The first preliminary experiment was designed to explore the impact of the visual appearance of a virtual human in comparison with the behavioural fidelity (gaze) of the virtual human on participants' perceived quality of communication in a shared VE. The hypothesis behind the experiment was that behavioural fidelity would far outweigh the importance of visual appearance.

The second preliminary experiment was designed to explore the importance of consistency with regards to the visual realism of different elements within a virtual scene. However, the relevance of the experiment to this thesis was its subsidiary goal which was to investigate the impact of interacting with virtual humans with limited expression as gauged through participant interviews.

1.2.2 Main experiments

The first main experiment was designed to investigate the importance of postural cues in the communication of affect in IVEs. The virtual humans in the experiment were designed to portray two affective states (Anger and Sadness) through postures and facial expressions in order to investigate the importance of congruent behavioural cues in full-body virtual humans. A combination of subjective and objective participant responses were collected with the view to determine a) how accurately participants recognised the underlying emotional state and b) to what extent participants respond appropriately.

The second main experiment was designed to investigate the importance of body movement, in addition to key postural cues, on the communication of affect in IVEs. The virtual humans in the experiments were designed to portray the same two affective states. The influence of facial expressions were eliminated through the use of a mask in order to ensure that any participant responses were due to the affective kinesic cues of the virtual human. Once again, participant responses were collected to determine the extent to which the participant recognised the underlying emotional state of the virtual human and the extent to which the participant responded to it.

1.3 Terminology

Conventionally, virtual humans are categorised, in accordance to *agency*, into two subsets: avatars and agents. Schroeder (2002) defines agency as the extent to which a virtual human is perceived by individuals to be a representation of other individuals in the physical world. The categorisation is based purely on whether the virtual human represents a participant in a VE or is a scripted artificial creation. Accordingly, agents would occupy the lower end of the agency spectrum while avatars occupy the higher end. An *avatar* refers to a virtual human used to represent an individual to others in a shared virtual environment (VE) while an *agent* refers to a virtual human with pre-scripted behaviours. Unlike the term avatar, "*agents*" have numerous definitions based on their characteristics, some of which have been summarised in Table I.1 in Appendix I.1. Within the scope of this thesis, the term *agent* is used to refer to a virtual human with a *perceived* internal psychological state that is reactive to a participant in the VE and perceived to be communicative by the participant.

Throughout the thesis, the term "*realism*" or "*realistic*" is used to define the visual or behavioural sophistication (fidelity) of a virtual human. The word has a number of connotations attached to it especially in conjunction with the term *believability*. For instance, when Disney use the word realistic, what they really mean is *exaggeration* or emphasis of a character trait (Thomas and Johnson, 1981). Realism in this thesis is meant to refer to a visual or behavioural aspect of the virtual environment (e.g. the agent) that is close to mimicking individuals in the physical world and capable of evoking realistic responses from the participant.

The visual realism of virtual humans can vary along three dimensions: anthropomorphism (Nowak and Biocca, 2003; Schroeder, 1996), photorealism (Schroeder, 1996) and truthfulness (Benford et al., 1995). Anthropomorphism refers to the extent to which the virtual human resembles a humanoid form while high photorealism refers to the extent to which the virtual human appears non-cartoonish with more visual details such as texture-mapped faces. Truthfulness is defined by Benford et al. (1995) to quantify the extent to which an avatar resembles its user. In terms of visual realism, all the virtual humans used within the scope of this thesis refers to humanoid representations with no regard to the dimension of truthfulness since this research explores the behavioural components of virtual humans. For instance, a visually realistic virtual human is one that has photographed faces from the physical world textured-mapped to the facial geometry of the virtual human.

A behaviourally realistic virtual human is one in which the animated behaviours are generated from real life behaviours displayed in the physical world. In defining computational models of behaviours, two terms are used repeatedly: *parametric* and *data-driven* (non-parametric). A parametric model simply refers to a behavioural models that is controlled through the use of a set of parameters. For instance, the parametric model for affective kinesic behaviours is controlled using a set of parameters including one that controls the speed of the body animation displayed by the virtual human. Kinesics refers to both postural cues and body movement. A data-driven model refers to a behavioural model which is built based on a database of pre-existing behaviour animations.

Finally, another term that is used in this thesis is '*presence*' and '*copresence*'[1]. Traditionally the effectiveness of a virtual reality experience is measured using standard post-experience questionnaires designed to ascertain the participant's subjective rating of their sense of *presence* in the environment. The effective representation of a virtual human is measured using the participant's sense of *copresence* with the virtual human. A participant's sense of presence is defined in many ways including one that refers to their sense of "being there" while an individual's sense of copresence refers to their sense of being in the company of others. In this thesis, an additional measure of presence is adapted from the approach taken in Sanchez-Vives and Slater (2005). In accordance to this definition, presence is taken as the extent to which participants act and respond to virtual sense data as if they were real, where 'response' is considered at many different levels ranging from physiological through to cognitive. Following this definition, copresence is taken to mean the extent to which participants respond to virtual humans as if the virtual human was real.

1.4 Scope of this thesis

The bulk of this thesis deals with trying to invoke different types of responses from participants in a VE using real-time interactive agents which mimic behaviours typically displayed by people in the physical world. To this end, the thesis explores two particular human behaviours, kinesics and gaze, with properties gathered from literature in social psychology. It presents a methodology that will be

[1] A brief overview of presence and copresence is given in Section 2.6 under Chapter 2

used to test the plausibility of these behaviours in controlled experiments. Parametric models of the behaviours are built and then tested through participant-virtual human interactions. The measures used to judge the effect of the modelled behaviours on participant responses in the VEs are also dealt with.

The virtual humans dealt with in this research are designed to resemble the human form (the high end of the anthropomorphism dimension) and are used in an IVE. Throughout the thesis the visual realism of the virtual humans vary from cartoon-like forms to semi-realistic forms with texture-mapped faces. One of the experiments, described in this thesis, was carried out with the use of avatars. The remaining three experiments deal with agents.

This thesis does not deal with the modelling of the internal emotional state of a virtual human. It does not deal with the relationship between personalities and emotions nor does it focus on the various number of behavioural animation techniques. The research presented in this thesis deals with the representation of internal states through the use of two behavioural cues: kinesics and gaze. Although, a parametric model of gaze behaviour is utilised to a great extent in the experiments, the main behavioural cue of interest in this thesis is kinesics (postural cues and body movement).

1.5 Contributions

A critical literature review of the challenges surrounding the creation of affective virtual humans is presented in Chapter 2 by drawing on the research conducted in multiple disciplines including social psychology and computer science.

- The review assumes that a virtual character representation is already available, and discusses a variety of models and methods that can be used to enhance the expressive capability of virtual humans. It also identifies gaps in existing models used to build expression in virtual humans. For instance, the review suggest that although there are specific behavioural attributes associated with some affective states, mapping those behavioural attributes to the affective state, in an interactive full-body virtual human, is not straightforward. Further analysis of the review revealed a gap in the knowledge regarding the expression of affective states through kinesics, in comparison to other modalities, such as facial expression, which have been more widely and systematically researched.

This thesis makes technical and empirical contributions to research into the creation of expressive virtual humans and understanding participant responses to virtual humans in immersive virtual environments. The research focused on two specific behavioural cues: *kinesic* and *gaze* behaviour. In both investigations, a parametric behavioural model was built for use in virtual humans. The gaze behaviour model was evaluated through an experiment designed to investigate the importance of behavioural fidelity in virtual humans (Chapter 4). The model of affective kinesics was investigate with respect to the communication of two affective states (Anger and Sadness) in two experiments (Chapters 5 and 6).

- The parametric model of affective kinesics is presented in Chapter 6. The model was used to portray two underlying emotional states in virtual humans through postural cues and the quality of body movement.

- The parametric model of gaze behaviour for virtual humans used in dyadic face-to-face communication within immersive virtual environments is presented in Chapter 4.

- The two parametric models were evaluated by observing participant responses to the virtual humans during interactions in IVEs. Findings from these evaluations suggested that participant experience more realistic responses to virtual humans if the virtual human is controlled by the parametric model of behaviours presented in this thesis. Findings relating to the impact of the parametric behaviour models on a variety of participant responses, including subjective evaluations and physiological responses, are discussed in Chapters 4, 5 and 6.

1.6 Structure

This thesis contains motivations, background and related work concerned with the communication of affect through nonverbal behaviours and the natural responses of others to the individual's emotional state.

Chapter 2 examines the existing state of art regarding the area of research. The chapter gives a review of existing behaviour generation systems and computational models of emotions in addition to a review of emotions from the social psychology literature. Challenges in creating affective behaviours in agents is discussed. For instance, one of the limitations to creating an expressive virtual human (agent) is the lack of data on the exact behavioural cues corresponding to an emotional state. A comprehensive review of emotions and the variety of behavioural cues used to express emotions is discussed. Parts of this chapter has been reported in Vinayagamoorthy et al. (2005) and Vinayagamoorthy et al. (2006b)

Chapter 3 addresses the methodology used in previous experiments that has led to the current state of the research and its adaptation in the design of the experiments reported in this thesis. It discusses the use of subjective measures (questionnaires and interviews), behavioural measures and physiological response measures. It also discusses the use of both quantitative and qualitative methods to triangulate explanations for statistical findings.

Chapter 4 presents two preliminary experiments which focused on exploring the impact of building realistic behaviour animations into virtual humans on participant responses. Findings from the first preliminary experiment has been reported in Garau et al. (2003) while the parametric behaviour model associated with the experiment was reported in Vinayagamoorthy et al. (2004b). Parts of the findings from the second preliminary experiment have been reported in Vinayagamoorthy et al. (2004a) and Brogni et al. (2006).

Chapter 5 presents an experiment designed to explore the role of postural cues in the communication of two affective states: Anger and Sadness. Part of the findings from this experiment has been reported in Vinayagamoorthy et al. (2006a). In Chapter 6, a potential method to evaluate the successful creation of a virtual human is presented with another experiment designed to investigate the role of body movement in the communication of affect.

Finally, Chapter 7 draws a conclusion on the work presented in the thesis and gives a summary of possible future direction in the research.

Chapter 2

Background

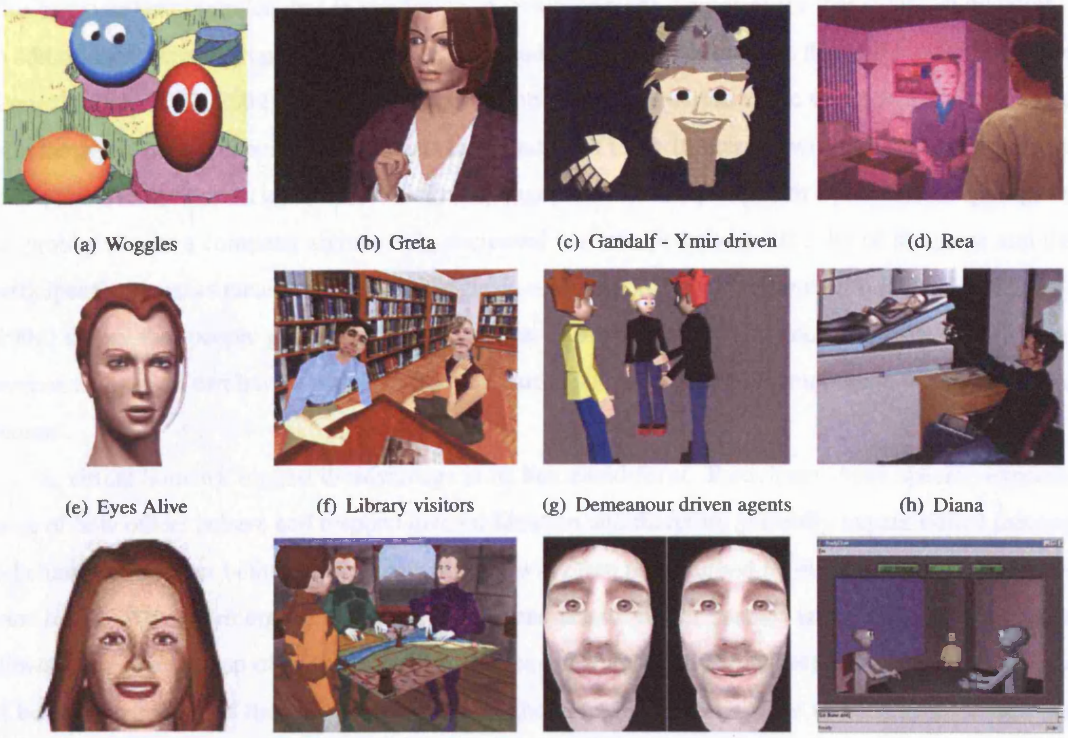
This chapter focuses on the providing background to the challenges met in building *affective* virtual humans. There are four parts to this chapter. The first part deals with the potential benefits of building expression into virtual humans (Section 2.1).

A critical review of the literature suggests that there are two distinct but interrelated research themes concerning the creation of an affective virtual human: the *generation* and the *representation* of an internal emotional state. The first research theme deals with the generation of a virtual human's internal emotional state. The second part of this chapter, Section 2.3, discusses some computational models of emotions used in this generation process. These models play an important role in simulating internal emotional states, in virtual humans, for applications that require long-term or repeated interactions with the participant. The second research theme deals with the representation of the generated emotional state through the use of different behavioural cues (modalities). The third part of the chapter deals with this research theme in two sections. The different functions on nonverbal behaviours, in the physical world, is reviewed in Section 2.2, however, Section 2.4 focuses on the different behavioural cues used to portray emotional states. There are many behavioural cues available to an individual in the physical world, however, the role of certain behaviours, such as kinesics, have remained obscure in comparison to others such as gaze and facial expressions. Despite the extensive research into expressive animation techniques, there is very little research into building parametric models of affective kinesic behaviour even though kinesics can play a significant role in the communication of affect in full-body virtual humans. In keeping with the scope of this thesis, Section 2.4 focuses on two behavioural cues: kinesics and gaze. Due to the importance of gaze behaviour in communication, a parametric model of gaze behaviour was developed and used throughout most of the experiments presented in this thesis. Section 2.4.2 discusses the role of gaze in dyadic communication. The focal point of this thesis is kinesics behaviour, in particular the role of posture and body movement in the communication of affect, which is discussed in Section 2.4.3.

The final part of this chapter focuses on responses to viewing expressions of emotional states in the physical world. Section 2.5 presents a set of participant responses, such as stress and anxiety, that could be invoked by affective virtual humans with particular focus on one emotional state: *Anger*, while Section 2.6 deals with the concept of presence as a means to measuring the success of virtual humans in

IVEs. Finally, Section 2.7 gives a summary of the chapter and how this thesis fits with existing work on building expression into virtual humans.

2.1 Benefits of building expression into virtual humans



(a) Woggles (b) Greta (c) Gandalf - Ymir driven (d) Rea

(e) Eyes Alive (f) Library visitors (g) Demeanour driven agents (h) Diana

(i) Happy agent, MIRAlab (j) Spark (k) DER driven agent (l) Bodychat

Figure 2.1: Examples of some expressive virtual humans.

The main benefit virtual humans offer in human-computer interaction is their ability to invoke the participant's automatic responses to the human form and behaviour. Evaluative studies suggest that virtual humans can elicit the appropriate and sometimes surprising responses from participants. It has been suggested that avatars with minimal behaviours can increase a participant's perceived quality of communication (Garau et al., 2001) while increased agent responsiveness can have a significant effect on participant's social responses (Garau et al., 2005). There is also evidence that participants who are prone to paranoia in the physical world are more likely to become anxious in response to virtual humans in essentially neutral contexts such as those depicted in Figure 2.1(f) (Freeman et al., 2003). This supports the premise that increasing behavioural fidelity and responsiveness in agents even on a simple level can have a significant impact on individuals in VEs.

Other studies have reported that the simulation of inferred behaviours in an agent can greatly affect an individual's experience in VEs. This effect is particularly true of responses and situations that are specifically designed to have a strong emotional content. For instance, individuals respond to praise, criticism, personalities and social responses from a computer or a virtual human as if it were real (Nowak and Biocca, 2003; Pertaub et al., 2002; Prendinger et al., 2005). Pertaub et al. (2002) found that participant's rating of their own performance at public speaking was significantly affected by whether they were given positive or negative nonverbal feedback from an audience made of completely scripted virtual humans. The audiences were made up of either positively inclined, negative inclined or emotionally neutral static agents. Participants who gave the presentation to the negative audience developed greater anxiety. This has significant implications in the design of virtual humans employed for therapeutic applications. In addition, participants reported that the negative audience was more realistic than the positive audience (Pertaub et al., 2001, 2002). Finally, the participants' self-rated performance was positively correlated with the perceived *good* mood of the agents independent of time. In keeping with this result, Prendinger et al. (2005) found that an agent whose body language expressed empathy with a person, and apologised for problems with a computer significantly decreased both the perceived difficulty of the game and the participant's stress, as measured via physiological responses. This is consistent with Reeves and Nass's (1996) theory that people generally treat computers as social actors. This means that even a minimal level of expression can have a positive effect on participant responses to virtual humans under a social context.

A virtual human's biggest disadvantage is its humanoid form. Participants have specific expectations of how others behave and respond given a situation and therefore generally expect virtual humans to behave in a manner befitting its appearance and will often be disturbed by discrepancies in its behaviour. In fact, it has been argued that the visual appearance of virtual humans is only important in that it allows for the generation of social behaviour (Bailenson and Blascovich, 2004) and that the importance of behaviour fidelity of the avatar far outweighs the visual realism in some applications (Blascovich et al., 2002). In a study on collaboration in a shared VE, groups of three participants were represented by avatars and asked to solve a series of puzzles in a virtual room together (Slater et al., 2000; Steed et al., 1999; Tromp et al., 1998). Although, all three participants were represented by simple avatars with limited behavioural capabilities, one of the avatars (Green) was made visually more realistic than the other two. Post-experimental discussions revealed that the differences in the green avatar's representation had an effect on the relationship formed during the collaboration. The increased visual realism raised the other participant's expectation about Green's capabilities which were not fulfilled:

> *The Red subject found that the Green avatar was 'scary', 'like a zombie' because there was a conflict between its greater visual realism but lack of bodily movement. Blue on the other hand was more 'cartoon-like' and therefore it was easier to understand the fact that it was not 'functional'.*
>
> *- Anecdote reported by Slater and Steed (2002)*

Nowak and Biocca (2003) conducted a study to examine the influence of anthropomorphism and perceived agency on the subjective presence responses of individuals (Table 2.1). Participants were required to take part in a very limited interaction with a humanoid representation on a screen. The perceived agency of the representation was induced by telling individuals that the representation was either an avatar or an agent depending on the condition and the anthropomorphism factor of the representation was varied by using either a humanoid face, an abstract image (only mouth and eyes) or no image at all. Interestingly Nowak and Biocca (2003) found that participants reported a higher sense of copresence while interacting with a less humanoid representation. They argued that this was inline with the speculation that individuals interacting with a highly anthropomorphic virtual human had higher expectations with regards to the ability of the virtual human which were not met in the study.

	Perceived as an avatar	**Perceived as an agent**
High Anthropomorphism	low copresence	low copresence
Low Anthropomorphism	high copresence	high copresence
No representation (*control*)	low copresence	low copresence

Table 2.1: Conditions and some results from Nowak and Biocca (2003)

One consensus that is emerging is that the virtual human's visual appearance and behavioural fidelity need to be consistent (Schroeder, 2002) and that an accurate visual likeness to the human form is not necessarily the primary factor when trying to portray the "*realism*" of a convincing virtual human. This is reflected in the results reported by Nowak and Biocca (2003). The interpersonal communication of emotions, interpersonal attitudes, personality traits within individuals is integral to regulating the communicative and behavioural ability of virtual humans (Argyle, 1969; Picard, 1997; Gratch et al., 2002). Building virtual humans controlled by affective behavioural models can greatly enhance the types of participant responses elicited. In effect, nonverbal behaviours are an important tool through which the virtual human could convey a psychological state. However, as already stated creating behaviour models in virtual humans is not straightforward when the application requires run-time interaction or face-to-face communication.

2.2 Functions of nonverbal communication

Generally face-to-face communication channels can be divided into two distinct but interrelated categories: *verbal* and *nonverbal*. The verbal communication of an individual's psychological state is undertaken using both literal (*'I am irked, angry or outraged'*) and figurative (*'blowing a gasket'*) statements. In addition to content, every verbal message contains an insight into the psychological state of the individual and the relationship between the individuals in the conversation (Duck, 1998; Watzlawick et al., 1968). Nonverbal behavioural changes give a tone to face-to-face communication, accent it and sometime even overrides the verbal part of the communication (Argyle and Trower, 1980). Even though verbal and nonverbal content might not always indicate the same message, what they convey is almost always compatible (Gratch et al., 2002). However, although it is important to achieve synchronicity between verbal and nonverbal communication, this thesis does not deal with verbal communication.

Even simple behaviours like the *smile* depend on many factors including *culture, interpersonal relationship* and *context*. Some cultures use the *'smile'* as a cover during times of embarrassment. In addition to culture, the *'smile'*, normally associated with pleasure and a positive response, is sometimes generated in aversive situations such as when the context of the conversation is rooted in sarcasm (Darwin, 1872). The *smile* is also used when an individual is trying to put across a negative opinion in an amicable manner. A type of *smile* displayed in one context can be attributed a completely different meaning within another context. Knowledge of the cause and context within which it is expressed can greatly expand its interpretation (Argyle and Trower, 1980; Scheflen, 1964). For instance, if an unfriendly message is delivered with a smiling facial expression, the message is taken to be friendly (Argyle and Trower, 1980) and Wiener et al. (1972) have argued that the individual listening in a conversation can use the *smile* in conjunction with eye contact to signal comprehension but an unwillingness to verbalise this. Smiles also function to make communication more efficient by providing the individual who is speaking with feedback on a number of levels simultaneously including level of understanding, agreement, response and involvement (Brunner, 1979).

Nonverbal communication is a highly complex but vital component of creating expressive virtual humans. Ekman (1965) suggested that nonverbal behaviours systematically change as a function of a gross modification in the quality of the relationship between individuals. This view is further supported by studies on the impact of praise or insult on the willingness of individuals to engage in mutual gaze (Exline and Winters, 1965). This suggests that in designing virtual humans for use in applications which require interpersonal communication, it is important to gain a thorough understanding of nonverbal behaviours and the meanings attributed to these behaviours. Additionally, it is important to take into account the context within which a virtual human is used and also be sensitive to other confounding variables such as emotional states, personality traits and cultural differences.

A great degree of controversy exists in the definition of which nonverbal behaviours can be regarded as nonverbal communication. In face-to-face communication, nonverbal behaviours can function to create a variety of different connotations to the same verbal contents of a message including: emphasis, contradiction, accents, and affect (Ekman, 1965). An early view put forward by Ekman and Friesen (1969) argued that only nonverbal behaviours which are intended to be communicative can be regarded as nonverbal communication. On the other hand, it has been argued that the criterion of intention to communicate is irrelevant and that all behaviour conveys information (Watzlawick et al., 1968). According to the later definition, all nonverbal behaviours are communication centred. For instance, individuals can pretend to express an emotional state they do not feel (e.g. an interview candidate who feels nervous but attempts to hide this through a rehearsed firm handshake) or unintentionally express an emotion they are trying to hide (e.g. an interview candidate whose words sound confident but whose posture reveals nervousness). Displaying behaviours without the conscious intention to communicate a psychological state, such as boredom, is termed emotional *leakage* and normally occur in body posture since individuals pay less attention to the control of posture than facial expressions (Bull, 1987). Wiener et al. (1972) make a distinction between behavioural responses (*signs*) and goal directed action (*communication*). Distinguishing between signs and communication has a lot to do with the awareness of both the individual expressing the message and the other (the individual interpreting it). While communication is always directed at another individual, a sign can either be due to another individual (e.g. involuntary laughter) or something that is completely unconnected to the other individual (e.g. making a hunched posture when cold). The difference between these three cases can be thought of in terms of how an agent could react to an individual in a VE. If an agent is unaware of the individual's presence, it will just produce signs which are not directed at the individual whereas if the agent is aware of the individual it will use communication. However, if it is trying to ignore the individual it will not communicate but some involuntary signs aimed at the individual might be 'leaked'. According to (Wiener et al., 1972), it may be that all behaviour is potentially informative but this is something that has to be demonstrated not assumed.

Individual	**Other**	**Examples**
Aware	Aware	Verbal communication and some gestures
Mostly unaware	Mostly unaware	Most nonverbal *communication*
Unaware	Unaware but has an effect	Pupil dilation, gaze behaviour and nonverbal *signs*
Aware	Unaware	Spatial behaviour (Proxemics)
Unaware	Aware	Other is a psychoanalyst of e.g. body posture and quality of body movement

Table 2.2: Distinguishing types of communicative nonverbal behaviours

The view held within the scope of this thesis is similar to that of Bull (1987) and Scheflen (1964), in that neither intention nor awareness of specific behavioural cues is necessary for regarding nonverbal communication to have taken place between individuals especially in a face-to-face situation with regard to bodily communication. The process of human face-to-face communication contains superfluous behaviours which may not directly contribute to the comprehensibility of an incident (Birdwhistell, 1971). They are not only essential to a conversation but adds to the richness and versatility of the conversation (Argyle, 1998; Argyle and Trower, 1980). This apparent redundancy often serves to reinforce the verbal contents of the message. The view adopted in this thesis is that nonverbal behaviours exist to reduce the ambiguity of a message and therefore nonverbal communication may take place between individuals without any conscious intention to communicate (Scheflen, 1964). In some cases, communication occurs against the conscious intentions of the individual (Bull, 1987) whereas in other cases an individual expresses nonverbal behaviours while trying to communicate even though the conversational partner is on the other end of a telephone call and could not possibly *see* the expressions (Argyle and Trower, 1980). In virtual humans, behavioural cues are generally designed with the intention to communicate a specific psychological state, however, the participant's interpretation of the virtual human's behaviour might include more than the intended psychological state.

Nonverbal communication has many different types of functions. The first type is the expression of an individual's psychological state such as:

- The provision of information about the individual's feelings or attitudes to others in an interaction (Mehrabian and Firar, 1969). Nonverbal behaviours function as cues in the expression and perceived intensity of the emotion (Patterson et al., 1986). These are often called *Affect Displays* (Figures 2.1(i) and 2.1(k)).

- Nonverbal behavioural cues are often used to either project a specific personality trait such as dominance or unconsciously provide cues to the individual's persona (Mehrabian, 1971; Mehrabian and Firar, 1969) (Figure 2.1(g)).

- The information can also be about someone's cognitive state. For example, looking away or expressing concentration can indicate reflection, while other facial expressions can show comprehension or confusion (Poggi and Pelachaud, 2002; Pelachaud and Poggi, 2002) (Figure 2.1(b)).

- *Adaptors* are self-touching behaviours that provide insight into an individual's attitude and anxiety level since they are the least consciously monitored behaviours. They are mainly used unconsciously (e.g. wringing your wrists) and therefore provide a rich source of involuntary information about the individual.

Of this type the first, emotions, is the most relevant to this thesis and possibly the most important. It is discussed in detail under Section 2.3. As well as representing an individual's internal psychological state, nonverbal expressions can also given an insight into the relationship between two or more individuals. This second type of behaviour can be further divided into two: behaviours that convey information within a conversation and behaviours linked to interpersonal relationships. Typically the functions of the first sub-type of nonverbal behaviours, as identified through conversation, includes regulators, emblems and illustrators (Argyle, 1998). However, these are beyond the scope of this thesis. The second sub-type of behaviour gives information about long term interpersonal relationships between individuals including expression of intimacy, emotional closeness and self-presentation (Duck, 1998).

Condition	Feedback
Content only	Answering questions and executing commands
Content with envelope	Included behaviours such as gaze (aversion and lifting eyebrows), manual beat gesture, and head movements, tapping fingers, attention directing gestures such as turning to the participant when listening and towards task when executing commands in addition to answering questions and executing commands
Content with emotional	Includes smiles when addressed by the participant and looks of puzzlement when asked sometime incomprehensible in addition to answering questions and executing commands

Table 2.3: Conditions of Thórisson and Cassell's (1999) envelope vs. emotional feedback study

It is very important to understand the various functions of nonverbal behaviours and how they interact when implementing an expressive virtual human. For instance, Thórisson and Cassell (1999) conducted a study comparing an agent (Gandalf - Figure 2.1(c)) that was only capable of conversation-driven behaviour (envelope) to one that was only capable of emotional expression. They found that people communicated better with the agent with conversational behaviour. This shows that it is important not to focus solely on one expressive feature such as emotion without considering certain communicative functions of nonverbal behaviour. However, this result should not be taken to mean that emotional expression is irrelevant. Firstly, a glance at Table 2.3 is sufficient to notice that the behaviours implemented to support the envelope feedback condition far outweighs the emotional condition. The interaction provided in the emotional feedback condition included smiling briefly when a participant addresses an agent and then speaking. This might lead individuals to judge the agent negatively due to the inconsistencies of the emotional expressiveness. A far better test would have been to include a fourth condition with both emotional and envelope feedback. This would have answered the additional impact of emotional expressiveness with the given context. Secondly, the effects of gaze and the importance of turn taking behaviour are not unknown. It is integral to any conversational task (Argyle, 1969). The communication of affect especially through gaze is only complimentary to the management of the conversation especially given the context. Secondly, the study involved talking about an emotionally neutral subject (astronomy). In an emotionally heated negotiation, participants may find the lack of emotional expression problematic or unrealistic. A possible conclusion is that some basic conversational behaviour is fundamental for virtual humans but that many other types of behaviour, such as emotional, are also important. Another conclusion is that what forms of expression are needed is highly dependent on the context in which a virtual human is used. For instance, emotional expression and building interpersonal relationships might not be important in a banking setting, but both are vital for a virtual human used in therapeutic applications.

This thesis focuses on the use of nonverbal behaviours as *Affect displays* and representation of internal psychological states. The equally important conversation management role of nonverbal behaviours is not explored in full though the significance of gaze in face-to-face communication is explored in a preliminary experiment (Section 4.1). Sections 2.3 attempt to summarise relevant theories and categorise existing computational models of generating emotional states in conjunction with factors such as personality. These models are often used to simulate emotional states in interactive virtual humans over an extended period of time where a variety of expression is important to add depth to the virtual human's behaviour. The manner in which the internal states of the virtual human could be expressed through behaviour is discussed in Section 2.4.

2.3 Emotions

Everyone knows what an emotion is; until asked to give a definition.

- Fehr and Russell (1984)

Emotions are loosely regarded as a reaction to personally significant events where the reaction may include physiological arousal, changes in cognitive processes, behavioural expression, action tendencies and subjective labelling of these feelings (Kleinginna and Kleinginna, 1981). Disney and other cartoonists have maintained that expressing emotional states is a necessary substrate for producing plausible characters (Thomas and Johnson, 1981). The character of Grumpy in Disney's version of "Snow White and the Seven Dwarfs" would not be the same without his regular expression of irritation and anger (Disney, 1937). Some researchers, like Picard (1997), argue that creating emotion is essential to creating perceived intelligence and reasoning in agents. The central idea is that emotions are always involved while thinking and should be simulated in virtual humans in order to express plausible behaviour.

The addition of an emotional dimension to a virtual human can have a significant effect on the interaction, however, modelling them is not straightforward. Evolutionary theorists argue that all emotions are innate and associated with a unique set of responses (Darwin, 1872; Ekman and Davidson, 1994; Izard, 1992; Schachter and Singer, 1962). This suggests that at least some behavioural cues can be used to portray emotional states in virtual humans. On the other hand, emotion theorists also argue that with the exception of some emotions (e.g. innate disgust and startle), most emotions are almost entirely learnt social constructions (Ortony and Turner, 1990). This means the behavioural cues used to express certain emotions are varied depending on individual differences, social context, culture etc. Any given emotion can motivate a variety of expressions and actions. In addition there is a lack of clear definition on the effects of changes in emotional states on behavioural cues. Nonverbal behavioural cues used to express an emotional state often combine a variety of different modalities including facial expressions, postural cues and body movement (Section 2.4).

2.3.1 The role of emotions

Emotions are responsible for generating a rapid and efficient response to important environmental stimuli which is useful for survival from an evolutionary point of view (Gratch and Marsella, 2005). In general, the primary function of emotions are to guide actions and provide information through facial, vocal and bodily expressions.

On a biological level, emotions prepare the body for actions like the fight or flight response in the face of oncoming threat (Frijda, 1988) such as during the invasion of an individual's personal space (Jeffrey, 1998). Emotions create the optimal physiological milieu to support the necessary behaviour in an emotionally charged event. On a cognitive level, emotions alter an individual's priorities thereby serving to allocate limited resources towards multiple plans and goals (Oatley and Johnson-Laird, 1987).In situations where hesitation could have adverse consequences, emotions function to allow rapid strategic planning (Gratch and Marsella, 2005). Emotions can arise out a set of a deliberate planning process which in turn can influence the decision making process. However, while allowing for rapid response, efficient social communication and adaptation, the affective system is also prone to errors. Once emotional situations escalate, guiding focus to the immediate and relevant goal makes individuals loose perspective thereby leading to irrationality. This is the main argument against incorporating emotional models into agents. It has been suggested that individuals develop coping strategies to manage their emotional states leading to the development of models simulating this relationship in virtual humans (Marsella and Gratch, 2002; Paiva et al., 2004).

Emotionally charged events are generally more memorable than unemotional ones, even more so for negative events (Thorson and Friestad, 1985). Emotional states can be seen as an important factor in retrieving specific memories and also as a useful method to index perceived memories in virtual humans. For instance El-Nasr et al. (2000) use a learning component to define the expectations of an agent using information on past events like the generation of a fear response in expectation of an undesirable upcoming event. Lim et al. (2005) use long term emotional memory and emotional tagging to influence the re-experiencing of events. Scheutz (2004) suggests a number of potential roles for emotions in agents including action selection (decision making), goal management, social regulation, learning, and memory control.

2.3.2 Emotions and moods

Mood represents the overall view of an individual's internal state. In the physical world, an individual's emotion guides actions while their moods function to shift not only cognitive content but also their processing mode (Ekman and Davidson, 1994; Isen et al., 1987). Isen et al. (1987) argued that positive moods facilitate cognitive flexibility resulting in more creative responses, more remote associations and an increase in the perception of relatedness among thoughts. Even mildly positive affective states in an individual profoundly effects the flexibility and efficiency of thinking and problem solving (Isen, 1987). Other than functionality, an affective state is differentiated as an emotion or a mood based on three other criteria: *temporal, expression* and *cause*. Emotions are brief lasting for a matter of seconds or at most minutes. Emotions are often associated with a facial expression (Ekman and Friesen, 1976) and have identifiable cause. Moods last for longer and are not associated with a specific expression or cause.

Modulating the actions caused by an emotion becomes difficult if it occurs during a mood causing cognitive instability. Emotions are phasic changes superimposed on moods which in turn can be thought of as the affective background. Despite these differences, emotions and moods are inextricably linked. Emotions can lead to particular moods, and moods can alter the probability that a particular emotion will be triggered and its intensity (Neumann et al., 2001; Ekman and Davidson, 1994). For instance, an individual in an irritable mood becomes angry more readily than usual and the resulting anger is more intense, decays more slowly and is more difficult to control. There is no research to determine if this is because the individual is in a continually low level of anger and readily provoked or because there is a difference in thresholds and related cognitive appraisals characterising the mood. Most existing systems represent moods as a low level of emotional arousal for a longer duration than emotions (Velásquez, 1997; Kshirsagar and Magnenat-Thalmann, 2002a; Rosis et al., 2003). Becker et al. (2004) developed a presentation agent (Max) that expressed a coherent portrayal of emotions over time including the dynamics of emotions and moods over time.

2.3.3 Intensity of emotions and emotional decay

The intensity of emotions is affected by a set of variables which include how important the event is to the individual, the level of unexpectedness associated with the event, the prevailing mood of the individual the individual's gender and arousal (Frijda, 1988). With the exception of gender, an increase in these factors intensifies the emotion while an increase in the period of time from the eliciting event results in emotional decay. The expression of emotional intensity and decay in a virtual human is generally implemented in some form (Velásquez, 1997; Gratch, 2000; El-Nasr et al., 2000; Prendinger and Ishizuka, 2001). Many computational models use emotional intensities as one of the ways to create emergent personalities in the agents (Paiva et al., 2004) while others use a model of personality in conjunction with emotions to create agent emotional states with different associated reactions to the same event (Prendinger and Ishizuka, 2001; Egges et al., 2002; Gebhard, 2005). For instance in Bates et al.'s (1994) Woggles (Figure 2.1(a)) and Rosis et al.'s (2003) Greta (Figure 2.1(b)), emotional intensity is assigned depending on the uncertainty in the agent's beliefs and the importance of achieving a goal.

Generally, the relationship between physical expression and emotional intensity is modelled in a linear fashion. This is in keeping with results reported by Hess et al. (1997) which indicate that the perceived intensity of the underlying emotion of morphed natural faces is linearly related to its actual physical intensity. In addition, Hess et al. (1997) found that the more intense the emotional expression, the more accurately the target emotion was recognised. This result goes some way in explaining the success of exaggerating expressions to enhance the plausibility of virtual characters (Thomas and Johnson, 1981). On the other hand, Bartneck and Reichenbach (2005) recently found a non-linear relationship between the physical intensity of synthetic faces and perceived intensity. The synthetic faces were produced by interpolating between a neutral and a face expressing an emotion of maximum intensity as defined by Ekman and Friesen (1976). Bartneck and Reichenbach (2005) reported similar results to Hess et al. (1997), in that the recognition accuracy of an emotional expression increased with physical intensity but *only* up to a certain point beyond which the accuracy levels do not vary significantly. Even though this suggests that the exaggeration of behaviours will only be worthwhile to a specific point, the emotional states of virtual humans are more likely to be recognised if an intensified physical expression is used. However, there is no research to indicate if participants will respond to a virtual human's physical expression of emotional states appropriately or if the participants' responses will be correlated with a more intense physical expression of emotion.

In addition, some interesting differences in the stimuli set used in both studies were uncovered. Bartneck and Reichenbach (2005) observed that the expression of anger was the more sensitively perceived emotion category suggesting that individuals are prone to detecting varying levels of anger more acutely than other emotions[1]. Hess et al. (1997) went further to report gender-related differences. Expressions of anger and disgust portrayed by males were rated as more angry and disgusted while expressions of happiness portrayed by females were rated as more happy. These results also support the premise that individuals are more sensitive to the physical expression of anger.

2.3.4 Personality

Moffat (1997) differentiates between emotions and personalities over two dimensions: *duration* and *focus*. Whereas emotions change over time, personalities remain more constant and are not specific to particular events. Personalities arise out of more indirect and long-term factors. An emotion is a brief, focused change in personality. Personality traits come into play when a virtual human is used in an application that is meant to create some sort of relationship with an individual or in cases where a group of virtual humans are placed in social setting. Personality represents the unique characteristics of an individual and plays a partial role in determining the manner in which the individual chooses to portray a particular emotional state.

In psychology, the aim is to represent and understand the human psyche. This is done through defining various dimensions to generalise possible personality traits amongst individuals and scale them in some way. Many of these models have been used to create personality in agents (Breese and Ball, 1998; Kshirsagar and Magnenat-Thalmann, 2002a; Gebhard, 2005). This approach to personality mod-

[1] See section 2.5.2.1 for a discussion on the subconscious response to threat-related stimuli

elling helps in designing virtual humans that have certain characteristics which partly determine their behaviour in a consistent and coherent manner. For instance, a friendly virtual human is designed to act in a friendly manner in any situation because of the traits in its personality. Isbister and Nass (2000) reported that participants were able to identify the personality of agents and preferred an overall consistency in agents regardless of whether the agent was matched to the personality of the individual or not. This suggests that agents with a friendly smile are expected to maintain friendly and open body postures. Granted this detracts away from modelling personality quirks as strong as Bugs Bunny (Reilly, 1997), however, it does given a starting point.

2.3.5 Models of emotions

The categorisation of emotions is as fuzzy as the definition of emotions. For instance, there is little doubt that Anger and Sadness are emotions but there is less agreement on the definition of moods (irritability), long-term states (love), dispositions (benevolence), motivational feelings (hunger), cognitive feelings (confusion) and calm states (satisfaction) (Brewer and Hewstone, 2004). Gratch and Marsella (2005) categorised approaches to modelling emotion into *communicative-driven* and *simulation-based* approaches.

Instead of modelling an internal emotional state, the communicative-driven systems focus on picking an appropriate display of perceived emotions based on end-goal of the agent. A lot of communicative-driven models use a model of basic emotions such as those defined by Ekman and Friesen (1978): *happiness*, *surprise*, *disgust*, *fear*, *sadness* and *anger*. This approach is well-suited for applications involving short-term interactions with participants. These applications include those designed to act as educational aids or e-commerce representative where the emotional range of the virtual agent is well-defined and relatively narrow. The communicative-driven approach is also sufficient to explore and map the different behavioural cues which portray a specific psychological state. This approach was used in the experiments discussed in Chapters 5 and 6.

On the other hand, the more sophisticated simulation-based systems attempt to simulate the appropriate internal state (emotion) which can then be used to choose the appropriate behavioural cues to portray the emotion. Most simulation-based systems are based on an appraisal theory of emotions such as the "Ortony, Clore and Collins" (OCC) model (Ortony et al., 1988) discussed in Section 2.3.5.2. Even though the end result is often aimed to be communicative, this approach could potentially afford more flexibility and variety in the expression. The simulation-based systems are more suited for applications during which participants interact with the virtual human for long periods of time. This approach allows the consequences of events on a variety of levels to be appraised in accordance to the goals, standards and attitudes of the agent before resulting in an emotional state. It works well for applications requiring decision-planning in a group of virtual humans (agents) or in circumstances requiring a human-virtual human interaction over a prolonged period of time.

2.3.5.1 Models based on basic emotions

The concept of basic (or pure) emotions was made famous by Ekman and Friesen (1978) and is commonly applied using morph targets to simulate emotional expressions as typified by Hess et al. (1997).

Each basic model of emotions is associated with a different set of emotional labels (Ortony and Turner, 1990). The six basic emotions as presented by Ekman (1982) were defined based on their association with a set of facial expressions (Ekman and Friesen, 1976). Due to it's simplicity, even in cases where a more complex model is used to generate an internal state, a basic model is used to communicate the internal state using the most appropriate physical expression mapped (Kshirsagar and Magnenat-Thalmann, 2002a). The most commonly used basic model is the one presented by (Ekman, 1982).

Ekman's (1982) complete set has been used in a number of early systems such as Velásquez's (1997) model of emotions (Cathexis) which was implemented on a 2D baby face - Simón (Velásquez, 1998). Each basic emotion had a family of related affective states in order to implement emotional intensity; for instance, fear was associated with fright, terror and panic. Cathexis also allowed for the modelling of emotion blending; for instance, grief was a mixture of sadness and anger or fear depending on context (Velásquez, 1997). Other attributes included in the model were emotional decay and moods. Mood was differentiated from emotion through a lower level of arousal and longer duration. The emotional state that was finally output from the model was then used to select an appropriate behaviour from a set of possible pre-defined actions (Velásquez, 1997). Currently Ekman's (1982) set of basic emotions (or a sub-set) is often utilised to categorise the end-effect of expressing emotional states after the internal emotion is appraised through more complex systems. These include Kshirsagar and Magnenat-Thalmann's (2002b) and André et al.'s (1999) models of emotion and personality, Ushida et al.'s (1998) emotional model based on fuzzy inference rules, Rosis et al.'s (2003) Greta (Figure 2.1(b)), and Tanguy et al.'s (2005) Dynamic Emotion Representation (DER) model (Figure 2.1(k)).

The most noticeable restriction in Ekman's (1982) set is the imbalance between categorising negative and positive affect. In order to overcome the unbalanced nature of Ekman's (1982) set of basic emotions, El-Nasr et al. (2000) created a set of emotions for evaluations of their Fuzzy Logic Adaptive Model of Emotions (FLAME) on a synthetic baby face, which included sadness, anger, joy, fear, laughter, pain, thirst and tired. Another restriction in the Ekman's (1982) set (as well as other basic emotion models (Ortony and Turner, 1990)), is the lack of sufficient labels to represent a rich set of emotional expression. Rosis et al. (2003) get over this by using more than two models of emotions: the Ekman's (1982) set, a sub-set of emotion labels from the OCC model (Ortony et al., 1988), and embarrassment and shame (Keltner and Buswell, 1996).

Schachter and Singer (1962) reported that participants were more susceptible to the mood of a confederate when they had no other explanation for an increased psychological state of arousal. This suggests that the context of a situation plays an active role in the emotional state felt by individuals. This is in keeping with results reported by Ushida et al. (1998) where participants reported more perceived emotional states than the actual six that were represented in very simple agents. Several personalities and motivations were attributed to the agents including the basic survival type motivations (thirst, appetite, and feeling good) and more complex social motivations (defending territory and communication). Freeman et al. (2003) reported similar results in which participants attribute sentience and had feelings of paranoia towards virtual humans (Figure 2.1(f)). This implies that for some applications involving expressive virtual humans, simple emotion models might suffice in producing plausible behaviour. It also re-enforces the theory that the same behavioural cues might be interpreted differently by participants depending on their understanding of the context in which the virtual human is presented.

In addition to Ekman's (1982) model of basic emotions, there are a number of other models (Ortony and Turner, 1990) such as Plutchik's (1980) model which contains four pairs of *opposites*: joy and sadness, acceptance and disgust, fear and anger, surprise and anticipation. Plutchik's (1980) theory is more balanced than Ekman's (1982) set of basic emotions, allows for emotional blends and varying emotional intensities (rage is more intense than anger). Albrecht et al. (2005) uses an emotion model, based on the "emotional wheel" described by Plutchik (1980). In this model, the emotional space is represented by a disk defined by two dimensions: activation and evaluation. Similarity between two emotions is proportional to the angle that separates their positions on the wheel. The emotional wheel model is used by some facial animation systems including those presented in Kurlander et al. (1996) and Raouzaiou et al. (2003).

2.3.5.2 Models based on appraisal theories

Ekman's set	**The OCC set**
Joy	Happy-for, Gloating, Joy, Pride, Admiration, Love, Hope, Satisfaction, Relief, Gratification, and Gratitude
Sadness	Resentment, Pity, Distress, Shame, Disappointment, and Remorse
Anger	Anger, Reproach, and Hate
Fear	Fear, and Fear-confirmed
Surprise	
Disgust	

Table 2.4: Basic emotions vs. OCC emotions

The most simulation-based approach to modelling emotions is to view the emotions as reactions which result from appraisals/assesments of events and objects in correspondence to goals (and probabilities of achieving them), beliefs, risks and attitudes. Appraisals can be basic sensory-information processing, can involve rapid and automatic cognitive processes or a much slower cognitive process. Plutchik's (1980) model goes some way in forming such a chain but a more commonly used and comprehensive appraisal model is the OCC model (Ortony et al., 1988). The OCC model provides a rule based system for triggering 22 emotions, however, this has been judged to be too complex and cumbersome for modelling in a virtual human (Bartneck, 2002; Ortony, 2003). Ortony (2003) revised the emotional structure of the OCC model to 10 containers by eliminating all the branches relating to the concerns of other virtual humans. The revised set of containers consists five positive and five negative type of affective reactions based on goals (hope, fear), standards (pride, anger) and taste (like, dislike). The argument is that the slight reduction in realism is a justified tradeoff for some applications given that the revised emotional structure reduces the computational complexity of the model.

Bates et al. (1994) built one of the first emotional agent (Woggles, Figure 2.1(a)) system on an architecture called Tok which consisted of Hap (action selection) and Em (emotional model). Em generated emotions based on the success/failure of goals as appraised by Hap. Em was built based on the OCC model and emotion intensities were generated based on the importance level of the goal. For instance, hope and fear in agents are the result of the belief that a goal had the chance to succeed or fail. In addition to generating emotions, Em also represented basic interpersonal relationships between the agents on a like-dislike dimension. For instance, the proximity of a disliked agent to an agent causes it to become angry (Bates et al., 1994). Ushida et al. (1998) also present an emotion model for simple spherical agents with a deliberative system based on the OCC but use a set of fuzzy inference roles which control the levels of seven emotional factors. These seven emotional factors control the emotional intensities of six basic emotional labels (Ekman, 1982). The emotional intensities are calculated using the emotional factors, time decay and other emotions. For instance, the higher the goal failure levels, the sadder the agent is. However, if the blameworthy levels are high as well then the agent gets angry. Ushida et al. (1998) reported that Happiness and Anger were easily recognised, followed by Sadness, Fear, Disgust and Surprise. Gratch (2000) built on the work done on Reilly's (1997) Em algorithm and further extended it to produce a generalised plan-based model of emotional reasoning in Émile. Émile allowed agents (in this case a pedagogical agent - Jack and Steve (Rickel and Lewis, 1998)) to appraise the emotional significance of events in relation to its own goals and the probability of achieving those goals. Émile was integrated in Marsella et al.'s (2000) IPD (Interactive Pedagogical Drama) system which, amongst other things, focused on the impact of emotional states (and intensities) on virtual human behaviour (Gratch and Marsella, 2001; Marsella and Gratch, 2001). Another way to represent the significance of goals and the agent's belief of achieving those goals is presented in Rosis et al.'s (2003) Greta. Greta used a representation of beliefs and goals to drive the generation of emotions and the decision to display the expression. The internal states of the agent were generated through the use of Dynamic Belief Networks and allowed for changes in emotional intensity with time, response delays, and blends (Rosis et al., 2003).

El-Nasr et al. (2000) suggested an approach to modelling the dynamic nature of emotions by simulating their effects on behaviour by using a learning process to activate blends of emotion that would affect and be affected by a number of factors including motivation. FLAME is based on an event-appraisal model which uses fuzzy rules set to map assessments of the impact of a certain event on pre-defined goals into an emotional intensity and state (El-Nasr et al., 2000). The generation of emotions were again defined using an event-appraisal model based on the OCC. The agent learns about the properties of different events through reinforcement learning and about the participant through a probabilistic model that keeps track of the participant's actions. For instance joy occurred out of occurrence of a desirable event. The emotional state of the agent changed in accordance to situation, time and experience. A decision making component gets an output from the emotional component to influence the agent's actions. Fuzzy rules are use to calculate the desirability of an event in accordance to the impact of the event on the agent's goals and the importance of achieving the goals. El-Nasr et al.'s (2000) FLAME also allows the agents to adapt it's behaviour to provide a wider variety of facial expressions. Evaluations of the agent, PETEEI (El-Nasr et al., 1999a), resulted in some emergent behaviours, for instance, an oscillation/confusion between extreme emotional states (El-Nasr et al., 1999b). However, the learning model significantly improved the plausibility of the affective behaviours expressed by the agents (El-Nasr et al., 1999a). Kesteren et al. (2000) follow the same principles as El-Nasr et al. (2000) and simulate natural emotional expression through the modelling of the OCC model using neural networks.

OCC-based appraisal models have been used in conjunction with other mechanism like coping (Gratch and Marsella, 2004) or social networks (Prendinger and Ishizuka, 2001). Gratch and Marsella (2004) focused on intensely stressful scenarios and therefore extended their unified model with the addition of a detailed model of problem-focused and emotion-focused coping to produce EMA (EMotion and Adaptation). In addition a simple personality model is used to allow the agent to choose which coping strategy it prefers to deal with a particular stressful situation. The coupling of the appraisal model with coping process models has led to some unexpected emergent coping behaviour (Marsella and Gratch, 2002). Similarly Paiva et al. (2004) created a 2D cartoon-like humanoid agents application (FearNot!) with an emotional model coupled with coping mechanism to evoke empathy in participants. The internal states of the agents are expressed through proximity, facial and body expressions. In a recreation of a bully-victim situation, young participants felt empathy towards the victim and anger towards the bully. Young participants, especially females, felt increased empathy if they perceived the agents to have followed their coping strategy advice (Hall et al., 2005). Prendinger and Ishizuka (2001) built their work on the premise that agent behaviour can not be generated by modelling internal states such as personalities, attitudes and emotions alone but has to be integrated social role awareness models. This allowed their agent to suppress the expression of an emotional state if it would result in the failure of a goal.

Few constructive additions have been made to the OCC model. Bartneck (2002) argued that a function of history should be incorporated into the model so that the same event occurring again would not result in the same emotional intensity. However, this is dealt with in the models using values to represent the desirability of specific goals/events. Picard (1997) and Bartneck (2002) point out that the OCC model is not designed for and therefore does not allow for interactions and dynamics between the different emotional states. Models based on the concept of basic emotions or the OCC model support emotional expression. The interactions and dynamics between emotional states is covered in models described in the next section which deal with mechanisms which elicit emotions as well.

2.3.5.3 Models based on primary, secondary and tertiary emotions

Sloman (2001a,b) categorise emotions into primary, secondary and tertiary emotions. The definition of primary emotions is similar to the definition of basic emotions in that they are defined as being innate. Primary emotions are produced by reactive mechanisms mapping external stimulus patterns to behaviours. For instance, the states that often elicit two major response patterns, *'fight or flight'*, are anger or fear respectively. Secondary emotions, such as hope, are learnt associations between recognised stimulus patterns generated primary emotions and analysed situations where these patterns occurred. Tertiary emotions arise from the interaction of emotions and other cognitive processes (e.g. motivation).

Scheutz et al. (2000) introduced the CogAff agent architecture which models agent's cognitive system into a reactive, deliberative (what-if processes) and meta-management (reflective process) layer. Primary emotions were triggered in the reactive layer, secondary emotions were triggered in the deliberative layer and tertiary emotions involve the meta-management layer. Evaluations suggested that in a simulated survival-type scenario, agents with reactive mechanisms and affective states could achieve the same goals more efficiently than agents with high-level deliberative mechanisms (Scheutz et al., 2000). Models based on Lazarus's proposed process involving primary appraisals, secondary appraisals and re-appraisals (Lazarus, 1991) allows for a much more dynamic representation of emotion process. Recently Tanguy et al. (2005) presented the Dynamic Emotion Representation (DER) model which represented changes over time in emotion intensities and the interactions between different emotions (Figure 2.1(k)). Emotional *impulses* produced by the mechanisms eliciting emotions, such as those based on the OCC model (Ortony et al., 1988), effect (and are effected by) the state of the DER. Primary emotions are used to trigger pre-organised behaviours that are associated to facial expressions as defined by Ekman (1982). Secondary emotions based on the Ekman's (1982) set are used to select facial signals corresponding to communicative functions. For instance, an individual with a high level of happiness might emphasise a word by raising his eyebrows where a person with a high intensity of anger might frown to achieve the same result. Figure 2.1(k) shows two types of smiles generated by the system depending on whether the agent is sad or happy. Tertiary emotions are used as filters on how emotional impulses effect primary emotions and how they change the intensities of secondary emotions. The DER model was built over models which elicit emotions from internal or external events such as those defined earlier.

2.3.5.4 Models of personality & emotions

A number of models focusing mainly on emotions tackle personalities by modelling emergent personalities. Ushida et al. (1998) model various personality types through the difference in emotional expression. For instance, the threshold levels for triggering an angry state in the agent is used to control the extent to which an agent is irritable. In Rosis et al.'s (2003) Greta (Figure 2.1(b)), personalities were implemented by varying the goal weights that change the importance agents attach to each goal. However, most applications which involve running a virtual human over a substantial period of time, call for a more robust personality model.

Generally, personality traits are used to set the threshold to generate emotional states and control the intensity of the emotion. Information about the agent's personality can influence the probability of choosing actions explicitly (Perlin and Goldberg, 1996) or introduce uncertainty in the generation of the virtual human's emotional state (Chittaro and Serra, 2004; Kshirsagar and Magnenat-Thalmann, 2002a). Chittaro and Serra (2004) present a goal-oriented approach to modelling agents. The personalities of the virtual humans are modelled through a probabilistic automata (Probabilistic Finite State Machines - PFSM) where behaviour sequences are chosen from an animation library (and sometimes modified) based on personality. Most systems simulating the internal states of agents include detailed models of both emotions and personality since the two are closely linked (Kshirsagar and Magnenat-Thalmann, 2002a; André et al., 1999). Two of the most prevalent personality models used in modelling individual characteristics are the five-factor model (FFM) (McCrae and John, 1992) and the PAD dimensional model (Mehrabian, 1980).

The relationship between personality and affective states is not emphasised in the FFM[2]. This explains the coupling of the FFM with the OCC model in many existing systems. Chittaro and Serra (2004) use the FFM of personality as input to their probabilistic automata based behaviour animation system. Breese and Ball (1998) modelled the current emotional state and long term personality trait in an extendable Bayesian Belief Network (BBN) model to create two-way causal relationships between the agent's internal states and physical expressions of these states. The model was used both as a means to perform both causal inferences and diagnostic reasoning while maintaining a level of uncertainty. Breese and Ball (1998) focused on two dimensions of emotional response: valence which represented the agent's level of happiness and arousal which represented the intensity level of the emotion. The two personality traits used in the model were dominance and friendliness based on the FFM as these were critical to modelling interpersonal relationship (Breese and Ball, 1998). More recently, Kshirsagar and Magnenat-Thalmann (2002b; 2002a) used BBN to model personality traits using the more well-rounded FFM coupled with a layer of mood. They argued that the model handled abstract concepts within a structured probabilistic framework and also handles uncertainty with respect to the generation of emotion. Personality was represented along a n-dimensional space (FFM) while emotions were represented as levels of arousal through an extended version of the OCC. Kshirsagar and Magnenat-Thalmann (2002a) added two other emotions, surprise and disgust, to the existing framework since these were not covered in original OCC model (Table 2.4). The mood of the agent was controlled through a probabilistic function of the agent's personality. For each emotional state probability, the system computes the probability of each possible mood state in relation to the personality of the agent and some thresholds. The mood state with the higher probability is selected and used to choose one of the emotional states and its associated facial expression. The overall emotional state of the agent depended on the emotional impulse caused by an event, the personality, the mood, time-linear emotional decay, and the previous level of emotional arousal of the agent (Egges et al., 2002). Egges et al. (2002) extended this model and linked it to a dialogue system (modelled using Finite State Machines) represented by a 3D face (Figure 2.1(i)). This model is further detailed as the PME model (Egges et al., 2004) and was also integrated with an idle motion synthesiser to create idle motions appropriate to the emotional state of a virtual human (Egges and Magnenat-Thalmann, 2005).

Similarly, André et al. (1999) presented an integrated model of emotions based on the OCC (Ortony et al., 1988) and personality based on the FFM (McCrae and John, 1992). The integrated model was used as a filter to constrain a decision process to control an agent's behaviour. Initially their model simulated agent behaviour based on 4 basic emotions (anger, fear, happy and sad) and 2 dimensions of personality (extraversion and agreeableness). André et al. (1999) then developed their model to include two affective information processing channels: reactive and deliberative. This is similar to the first two layers of Scheutz et al.'s (2000) CogAff architecture. Unlike Tanguy et al.'s (2005) DER model, the deliberative channel generated secondary emotions in accordance to the OCC (André et al., 1999).

[2]The FFM is sometimes referred to as the OCEAN model.

Gebhard (2005) presented ALMA which focused on the temporal variations in affective states. ALMA (a layered model of affect) simulated short, medium and long term affective states through modelled emotions, moods and personality respectively (Gebhard, 2005) based on the OCC model and FFM. Like Kshirsagar and Magnenat-Thalmann's (2002a) model, the personality traits of the agent was used to control the computation of the intensity levels of emotional states. Romano et al. (2005) model social knowledge in addition to personality (FFM) and emotions (OCC). The main disadvantage of using the OCC model in conjunction with the FFM model is the unclear mapping between the two.

Mehrabian's (1980; 1996b) three-dimensional Pleasure-Arousal-Dominance (PAD) model allows modelers to input some of the links between personality and emotions. Different emotions and personalities are viewed as a variations along these dimensions. For instance, a score of low pleasure, high arousal and high dominance would be interpreted as anger while a score of low pleasure, high arousal but low dominance would be interpreted as fear. Some emotion modelers have chosen to reduce the dimensions in the PAD model to just two: pleasure and arousal, following Russell and Fernandez-Dols's (1997) "circumplex" model of the facial affect space instead of Ekman and Friesen's (1975) Facial Action Coding System (FACS) model. However, there are studies that argue that two dimensions are insufficient to completely handle aspects of facial affect (Schiano et al., 2000). Becker et al. (2004) focus on the modelling of a coherent portrayal of emotions over time in an agent - Max. The emotional engine behind Max consisted of two components. One to simulate the dynamics of emotions and moods over time. The other component acted as an emotion categorisation model based on the PAD model (Mehrabian, 1996b). Gebhard's (2005) used the Mehrabian's (1980; 1996b) PAD model to model moods in ALMA instead of personality. Mehrabian's (1996a) mapping between the PAD model and the FFM model was used to define the agent's personality. Gebhard (2005) suggest future additions to the ALMA model to include changes to the agent's mood in accordance to the emotional intensity. Alternatively, there are other dimensional models which can be used to model agent personalities and determine virtual human behaviour such as the one presented in Lim et al. (2005).

In addition to personality, another significant determiner of affective behaviour is the social context and interpersonal relationship between individuals in a situation. These factors are beyond the scope of this thesis. Section 2.3 discussed existing computational models which can be used to define the generation of an internal emotional state in virtual humans. However, all these models are of no use in designing an affective virtual human unless the relationship between an internal emotional state and its associated behavioural cues is uncovered. Each emotional state is associated with a wide variety of reactions and behavioural inclinations which are collectively referred to as response tendencies by Ortony (2003). These responses range from the physiological changes to voluntary and involuntary behavioural actions. This thesis concentrates on the behavioural cues (*Kinesics*) associated with an emotional state (*Anger* and *Sadness*). The next section discusses the various modalities of expressions available to the individuals (and a virtual human) with particular focus on gaze behaviour in face-to-face communication (Section 2.4.2) and affective kinesics (Section 2.4.3).

2.4 Modalities of Expression

It is an interesting fact about humans that they are often able to predict with reasonable accuracy how other individuals will respond to and behave in certain kinds of situations. These predictions are rarely perfect, partly because when we make them, we generally have imperfect information, and partly because the people whose behaviour and responses we are predicting do not always respond in the same way in similar situations.

- Ortony (2003)

There are two main issues involved in determining the expressions of internal states. The first revolves around the question: is there a set of distinct behavioural cues for an emotional state caused by a specific stimuli? The second concerns identifying common behavioural cues which portray a specific emotional state even if it is caused by *different* stimuli. In other words, what are the primary/distinct behavioural cues used to communicate an emotional state and what are the secondary/superfluous behavioural cues used to communicate the emotional state? Planalp et al. (1996) investigated the variety of cues used to perceive emotions in naturally occurring situations. Majority of the participants (97%) reported using more than a single cue. On average 6 to 7 cues were used, 13 cues being the maximum number reported. Two thirds of the participants reported using vocal cues while over a half used facial, indirect verbal and context cues. In terms of accuracy, 84% of the participants correctly perceived a single emotion and 68% matched all emotions in cases where multiple emotions were felt. This suggests that individuals, in the physical world, use different modalities of expression in a combined and highly synchronised fashion in portraying an emotional state.

Individuals express various emotional states, interpersonal attitudes and conversational feedback using a range of behavioural cues including: facial expression, posture and body movement (*kinesics*), gestures, head orientation and eye gaze, vocal intonation, and linguistic expressions. The most important determinant of the tone of verbal communication is the emotional state of the speaker (Argyle, 1969). A louder voice is often used by more assertive individuals to emphasise a point and also by individuals trying to express strong emotions like anger, surprise and fear (Argyle and Trower, 1980). There are studies into producing emotional tones of voice synthetically (Cahn, 1990), research into producing expressions based on recognising affect from the individual's voice (Pelachaud et al., 1994) and a combination of both (Cosatto and Graf, 2000). This thesis focuses on the nonverbal (and non-vocal) aspects of emotional expression.

The aforementioned modalities of expressions also emerge from the communicative process as part of *communicative acts*. Communicative acts can be acted, symbolic or intentional, and arise due to the communicative process in contrast to emotional expressions which arise due to emotional events. Communicative acts are also fast due to their synchronisation with the speech, in comparison to emotional expressions which have their own time signatures. Cassell and colleagues have developed a number of virtual humans (Figures 2.1(d), 2.1(j) and 2.1(l)) that exhibit realistic conversational behaviour (Cassell et al., 1999; Vilhjálmsson and Cassell, 1998; Vilhjálmsson, 2005) while Rosis et al. (2003) present an integrated system which translates goals, belief, and emotions into communicative function tags through dynamic belief networks (Figure 2.1(b)). With the exception of gaze behaviour, the equally important communicative role of nonverbal behaviours is not explored in detail within this thesis.

The following sections give an overview of various nonverbal behaviours used in the physical world, their functions and how the behaviour modelling of existing virtual humans have been tackled. Gaze, with respect to conversation management, and Kinesics, with respect to the communication of affect, have been covered in more detail since these modalities are the most relevant to this thesis.

2.4.1 Facial Expression

One of the most expressive areas of the body is the face, capable of producing about twenty thousand different facial expressions (Birdwhistell, 1971). It is the area most closely observed during an interaction (Argyle, 1969). Facial expressions of an emotional state are readily recognised by others even in synthetic format (Bartneck and Reichenbach, 2005; Hess et al., 1997). Emotional intensities are also perceivable from an individual's face with a great deal of accuracy (Bartneck and Reichenbach, 2005; Ekman and Friesen, 1978; Hess et al., 1997). Due to this, most methodological research has focused on the role of facial expression in the communication of affect at the expense of posture and body movement[3]. Facial cues associated with emotional states of anger and sadness are listed in Table 2.5.

Emotion	**Observable facial cues**
Anger	brows lowered and drawn together
	vertical lines appear between brows
	lower lid is tensed and may be raised
	upper lid is tense and may be lowered due to brow's action
	eyes have a hard stare and may have a bulging appearance
	lips are either pressed firmly together with corners straight or down *OR*
	lips are open, tensed in a squarish shape
	nostrils may be dilated (could occur in sadness too)
Sadness	inner corners of eyebrows are drawn up
	skin below the eyebrow is triangulated, with inner corner up
	upper lid inner corner is raised
	corners of the lips are drawn or lip is trembling

Table 2.5: Facial cues to some emotional states

Ekman (1992) and Izard (1971) conducted studies across different cultures in which they showed pictures of facial displays of emotion and asked participants to associate each picture to an emotional label. From these studies, a limited set of universally recognised facial expressions were identified such as Ekman's (1982) six basic emotions. According to Ekman's (1982) approach, meanings can only be expressed through a full-face configuration. Even if the facial patterns could be described by its component (Action Units - AU) using the Facial Action Coding System (FACS) (Ekman and Friesen, 1975), each component does not have any meaning by itself. Smith and Scott (1997) argued against

[3]Body movement includes postural attributes, body orientation, body agitation (tenseness), speed of body movement and fluidity of body movement (jerkiness).

this approach due to the limited set of meaningful facial expressions it provides and suggest that each component in itself has a meaning. This allows for the combination of the each AU to form a more extended set of facial expressions.

The FACS itself is a very useful to represent facial expressions but it does not provide any information about the meanings. It was developed as a standard method to code facial movements from images or videos but now it is widely used in computer animation. Existing systems either display one set of the universally recognised facial expressions (Kshirsagar and Magnenat-Thalmann, 2002a; Velásquez, 1997; Paiva et al., 2004) or in addition they produce combinations of these facial expressions (Albrecht et al., 2005; Kurlander et al., 1996; Raouzaiou et al., 2003). Some systems use a basic emotional model, such as the emotional wheel described by Plutchik (1980), to choose a corresponding pre-defined facial expression (Albrecht et al., 2005). More complex systems use dynamic emotion representations to produce emotional expressions (Kshirsagar and Magnenat-Thalmann, 2002a; Tanguy et al., 2005). These types of systems represent the slow changes of emotional intensities and therefore provide a consistency mechanism to produce emotional expressions (Tanguy et al., 2005). EMOTE and FacEMOTE presented by Badler et al. (2002) are interesting solutions for changing the expressiveness of an agent. EMOTE is discussed in more detail in Section 2.4.3.2. There are other animation systems available for generating affective facial expressions. However, they are not discussed here, since they are beyond the scope of this thesis.

Even though facial expressions are the most accurately interpreted modality of expression, Carroll and Russell (1996) suggest that the face merely provides information relevant to an emotion but does not signal a specific emotion. They argue that the context within which a facial expression is shown determines the emotions perceived from the face in question. If in a study an emotion is called for as a forced-choice then individuals will pick the most plausible emotion to match the expression (Carroll and Russell, 1996). If individuals are shown a static emotionally expressive face with no clue as to what elicited it, they imagine a context (Planalp and Knie, 2002) and if the nonverbal behaviour is contradictory to the context, individuals will try to justify it (Argyle and Trower, 1980). In addition, facial expressions accompanied with incongruent bodily cues would result in ambiguity in the judgement of emotion (Argyle, 1998; Dittmann et al., 1965; de Gelder, 2006). de Gelder (2006) reported that participants took longer to recognise the emotional state of an individual with less accuracy when the individual portrayed incongruent behavioural cues (facial and postural). This suggests that although the face is the main modality used to display emotional states, in a full-body virtual human, facial animation may not accurately portray the intended emotional state.

2.4.2 Gaze

The mechanisms of gaze behaviour are well documented making it easy to model and study the impact of implementing gaze behaviour in virtual humans. For instance, during dyadic conversations, it has been observed that individuals listening, in a conversation, look at their conversational partner (the speaker) for longer periods of time and more often in comparison to the speaker (Argyle, 1969; Argyle and Cook, 1976). This aspect has been used in the parametric modelling of gaze behaviour in virtual humans

(Garau et al., 2001). In addition, the behaviour of gaze can be studied using an eye tracker leading to data-driven models (Lee et al., 2002). Gaze behaviour in virtual humans was used in one of the preliminary experiments discussed in this thesis to explore the impact of modelling virtual humans with minimal behavioural cues on participant responses.

The eyes are probably the most intense social signallers in the human face (Argyle and Trower, 1980; Langton et al., 2000). Langton et al. (2000) argue that humans have evolved neural mechanisms devoted to gaze processing in order to enable the rapid and automated procedures needed to cope with analysing other individuals' gaze movements and trigger reflexive shifts in their visual attention. A period of mutual gaze between individuals often acts as a signal to initiate an interaction causing the individuals to move closer and also to signify attention focus (Goffman, 1963; Kendon, 1967; Muirhead and Goldman, 1979). At the start of the conversation and during the end, the amount of gaze between the individuals is high, however, this levels off to reach a state of equilibrium during the conversation (Argyle and Ingham, 1972; Argyle and Cook, 1976). Kendon (1967) observed that the speaker will often avert their gaze when there is a hesitation during the discussion of cognitively difficult material. Ponder and Kennedy (1927) observed that the rate of blinking reduces when the individual is thinking or paying particular attention to objects in the environment. Turn-taking is also actively controlled by the gaze and mutual gaze behaviours exhibited by participants in the conversation. In fact, Richardson and Dale (2005) report results from an experiment using an eye-tracker which suggest that the strength of the relationship between the speaker's and listener's eye movements predict the degree to which the listener successfully comprehended the information given by the speaker. Torres et al. (1997) attempt to capture the behaviour of gaze with respect to the occurrence of a turn, theme and rheme[4]. Empirical analysis of transcripts of speech, gaze behaviour and head movements were carried out on videotaped dyadic conversations. The results indicated that turn-taking processes are very predictive of gaze behaviour. Torres et al.'s (1997) results also suggest that the information structure of the conversation partially determines gaze behaviour (Gandalf - Figure 2.1(c)). This implication could be of specific importance to agents used in learning or educational based environments. Also the careful modelling of gaze behaviour in an agent gathering information from an individual could portray the perception of an attentive agent.

Amongst these and other well known conversation management functions (Kleinke, 1986), gaze serves to indicate involvement in an interaction, attitude towards the other individual in the interaction (Ellsworth and Carlsmith, 1971; Exline and Winters, 1965), characteristic attitude and perceived trustworthiness (Duck, 1998). Fukayama et al. (2001; 2002) reported that regardless of visual fidelity, gaze behaviour in virtual humans can be used reliably to convey different impressions to participants. In a face-to-face interaction, individuals who stare too much can cause others to feel ill-at-ease or uncomfortable while those who do not look at their conversational partners enough make them feel bored or disapproved of (Argyle and Trower, 1980). Individuals reduce mutual gaze under negative situations. For instance mutual gaze is avoided when an individual is embarrassed (Argyle and Trower, 1980; Duck, 1998; Exline and Winters, 1965). Individuals seeking affiliation with other individuals, engage in more

[4] Cassell et al. (1994) divide utterances into a *theme* part and a *rheme* part. The theme is the part of the utterance that refers to the subject matter already known in the conversation. The rheme is the new information in the utterance.

mutual gaze or glances towards others in search of more involvement in the interaction from the other (Argyle and Trower, 1980). However, an increase in gaze is also a means of expressing dominance and is perceived as a threat. Exline and Winters (1965) conducted studies in a interview scenario in which individuals were placed under either positive, neutral or negative conditions. Results indicated that participants in the negative conditions significantly reduced looking at or engaging in mutual gaze with their aggressor. In the positive condition, there was a slight increase in mutual gaze (Exline and Winters, 1965). Gaze and mutual gaze is also affected in accordance to the emotional state of each individual in the conversation and the interpersonal relationships between them (Kendon, 1967). Ellsworth and Carlsmith (1971) conducted a study to test Kendon's (1967) suggestion that the amount of mutual gaze in a dyad has a significant effect on the participant's responses to both the situation and their conversational partner depending on the verbal content of the interaction. Their results confirmed Kendon's (1967) theory to some extent. Frequent eye contact in conjunction with positive verbal content resulted in positive participant responses whilst in the negative condition, frequent eye contact produced negative participant responses. However, the gaze direction was significant in that if the mutual gaze did not result in actual eye-to-eye contact, the results were not in accordance to the theory that mutual gaze intensified the emotional content of the conversation. Ellsworth and Carlsmith (1971) argued that not actually looking directly at the participant while delivering the verbal contents of the conversation might have induced feelings of being "sorry for" or given the impression of "tactfulness" thereby resulting in less negative participant responses to the negative condition. These parameters, once well defined, can be used to induce feelings of threat and dominance in virtual therapeutic applications.

However, despite being subtle, much like any other behavioural cue, gaze parameters are difficult to generalise since mutual gaze patterns are also governed through factors other than speech patterns such as seating position, proximity, age and gender to name a few. For instance, Muirhead and Goldman (1979) reported that mutual gaze occurred twice as much when individuals sat opposite each other than beside each other and that middle-aged individuals engaged in half the amount of mutual gaze in comparison to younger or older individuals. Individuals also tend to avoid mutual gaze, the closer they are spatially (Argyle and Cook, 1976). Gaze is also affected according to status and gender; females dyads of equal status exhibit more mutual gaze in comparison to dyads of males or opposite-sex (Argyle and Trower, 1980; Exline and Winters, 1965; Mulac et al., 1987). Generally, females engage in more mutual gaze than males and tend to engage in increased mutual gaze with their preferred partner (stronger perceived affiliation) in a three-way interaction (Exline and Winters, 1965). A review of the different factors that affect gaze behaviour is presented in Kleinke (1986).

Deng et al. (2005) distinguished two approaches taken to model realistic gaze behaviour models: parametric versus non-parametric (data-driven) models. To date there are two main data-driven models. Lee et al. (2002) observed gaze behaviour in individuals using an eye-tracking device and computed a gaze model based on the first order statistical analysis of the collected data. The model depicted various properties of the rapid motion with which the eye moves from one focused position to another (*saccades*). These properties included the time the eye spent at a focused position (*inter-saccade interval*), the time the eye took to get to a new position (*saccade duration*), the angle the eye rotated in order to get to the new position (*saccade magnitude*), the direction the eye moved in (*saccade direction*), and the non-uniform velocity it moved with (*saccade velocity*). The resulting model was simulated on an agent (3D face) (Figure 2.1(e)) and participants were asked to judge, amongst other things, how natural, friendly and lively the agent was. Overall the model outperformed two control conditions using a static and a random gaze model. Other than the gaze model, Lee et al. (2002) also reported a high correlation between the movement of the eyes and eyelids which could, theoretically, be incorporated into an integrated model. More recently Deng et al. (2003, 2005) focused on producing an eye movement model using non-parametric texture synthesis techniques. Deng et al.'s (2005) technique was based on observations, also noted by Lee et al. (2002), that gaze changes were associated with blinks and considered eye-gaze and aligned eye-blink motion together as an eye-motion-texture sample. These samples were then used to synthesise new eye motions. Evaluations were conducted using the gaze model on a 3D face against the model proposed by Lee et al. (2002) and a random gaze model. The random gaze model was the least favoured model while the model proposed by Deng et al. (2005) slightly outperformed the model proposed by Lee et al. (2002). Rehm and André (2005) conducted a study as a probe to obtain data in order to later develop an appropriate gaze model for virtual humans in a multi-way social interaction scenario involving two participants and an agent (Greta) in a game involving deception. Their analysis revealed that in general participants followed the gaze patterns observed in dyadic situations and maintained the speaker-listener relationship, however, they gazed significantly more towards Greta when listening to the virtual human. Rehm and André (2005) hypothesise that this effect could be due to the difficulties participants faced in interpreting deceptive cues in Greta, the novelty of interacting with Greta or it could be that participants felt more comfortable looking at a virtual human.

Due to the extensive literature available on factors that affect gaze behaviour, a number of agents have been programmed with parametric gaze behaviour models. Colburn et al. (2000) conducted a study to investigate the differences in an individual's gaze pattern when interacting with an avatar with different gaze behaviour models. The gaze behaviour model of the avatar was modelled using state machines triggered in correspondence to who was speaking and the time passed between states. The participants in the evaluative studies displayed gaze pattern more similar to those occurring during a real dyad, when there was an avatar with life-mimicking gaze behaviour (Colburn et al., 2000). This result strengthens the premise that modelling even minimal behavioural cues in virtual humans has the potential to elicit realistic participant responses. Garau et al. (2001) conducted a similar study in which pairs of participants were asked to carry out a negotiation task under four conditions: audio only (no gaze), random (random gaze and random head animation), inferred (inferred gaze and tracked head animation), and video tunnel[5]. The first and last conditions were implemented as control conditions. In the two avatar conditions, the individuals in the dyad were represented to each other by identical gender-matched above shoulder avatars. Unsurprisingly, the video tunnel was the most favoured medium while

[5]Video tunnels are designed to correct for camera offset by offering a face-on view of the conversation partner. A camera was placed behind a half-silvered mirror reflecting the image from an upturned monitor. While the participant looked at the reflected image, the hidden camera captured a face-on view that was sent to the conversation partner's monitor.

the random gaze condition was the least favoured. The most encouraging results were that the inferred condition was not significantly different to the video tunnel condition (Garau et al., 2001).

Similarly, Fukayama et al. (2001, 2002) implemented a simple gaze model using a two-state Markov model based on inter-saccadic interval, amount of mutual gaze maintained with the participant, and where the virtual human looked when it was not maintaining mutual gaze with the participant. Results from their study showed that virtual humans could convey a perceived impression through gaze alone. Like Colburn et al.'s (2000) objective results, Fukayama et al.'s (2001; 2002) subjective results not only indicated that individuals do respond to virtual humans with realistic responses but also that even simple behavioural models can elicit these responses. The models employed by Colburn et al. (2000), Garau et al. (2001) and Fukayama et al. (2001, 2002) were created using the simple guidelines observed by Argyle et al. (1972; 1973; 1976) and Kendon (1967). Colburn et al. (2000) suggested an addition of a transition time-multiplier to their dyadic model to simulate the effects of proximal influences on gaze behaviour when multiple participants are involved in an interaction.

The above systems model the surface gaze behaviour observed in a conversation, but not the underlying processes. Gaze has a significant function in directing an individual's attention to an object. There has been some work on linking gaze to an underlying model of attention for social interaction in an agent that detected other agents' gaze and used it to decide whether to start or finish a conversation (Peters, 2005; Peters et al., 2005). Poggi and Pelachaud (2002) formalised the meanings that can be communicated through any nonverbal behaviour, applied it to gaze behaviour and suggested a set of AUs which may correspond to these gaze sets. The geometrical properties tackled by Poggi and Pelachaud (2002) is a step towards formalising the animation of gaze behaviour and include the eyebrows, upper and lower eyelids, wrinkles and eyes. One of the functions defined by Poggi and Pelachaud (2002), deictic, was implemented in an agent called Mack by Nakano et al. (2003) as part of a model of grounding which included feedback and turn-taking mechanisms. Preliminary studies suggested that the model encouraged more nonverbal feedback in comparison to interactions between participants and an agent with no grounding model (Nakano et al., 2003). Integrating these parameters into existing gaze behaviour models might prove to be beneficial.

Deng et al. (2005) argue against the use of parametric model by stating that although the models themselves are compact and economical, they fail to capture important aspects in the data. However, data-driven models are highly customised to specific contexts and involved gathering data over some sort of training period. Data-driven approaches may be suitable for the design of virtual humans used in applications where the interactions between the virtual human and participant is short-lived such as in gaming characters. This calls for a trade-off depending on the type of application. Deng et al. (2005) argue against the use of complex analytical models like hidden Markov models to analyse captured data as the hidden states influencing gaze behaviours are not easily interpretable. While this maybe the case, evaluative studies have shown that parametric approaches based on careful psychological grounding can have a significant effect on the perceived quality of communication with a virtual human. Applications based on communication and social interaction could benefit from the rich knowledge and flexibility that could be gained through analysing data collected using tracking devices and building customisable parametric models. To date, there are no complete studies which compare the effects of parametric versus non-parametric models. Additionally there have been no gaze behaviour models which combine parametric approaches with the knowledge gained from non-parametric approaches. One of the preliminary experiments detailed in Chapter 4 investigates the importance of gaze behaviour in the perceived quality of communication. The gaze model utilised in the studies was a parametric model based on the data-driven model presented by Lee et al. (2002) and the parametric model used by Garau et al. (2001).

2.4.3 The Body

The body can be as expressive a medium of nonverbal communication as the face and eyes. It is particularly important at a distance. Montepare et al. (1999) argued that the first cues perceived by an individual when others are approaching to initiate a social interaction are embedded in body movement and gestures. This means that body movement can be an important tool of expression for virtual humans that are to be viewed at a distance. The same applies to low polygon virtual humans that may not have enough detail for complex facial expressions. Of course, bodily nonverbal communication is also important for high-polygon virtual humans in close up; without plausible body movement, virtual humans can seem stiff.

It is also well known that the body plays an active role in portraying personality traits such as dominance or interpersonal attitudes such as affiliation (Argyle, 1998). It is also generally accepted that body movements and postures are indicative of an emotional state, however, what is a matter of debate is whether the body is indicative of a specific emotional state (*quality*) or only indicative of the intensity of the emotional state (*quantity*). Ekman and Friesen (1967) suggested that the face conveys specific emotions, while the body conveys the degree of intensity of the emotion. However, a number of studies and observations have lead researchers to believe that the body can be a more dominant source of cues in the perception of emotions (Argyle, 1998; Bull, 1987; de Gelder, 2006; Dittmann et al., 1965; James, 1932; Mignault and Chaudhuri, 2003) and in some cases an equally accurate source of cues to an emotion as facial expressions (Walters and Walk, 1986, 1988). Bodily cues are especially thought to play more importance when facial expressions are not viewable, for instance, when the individual is at a distance (Walters and Walk, 1988). Walters and Walk (1988) report higher recognition accuracy in perceiving happiness, anger and sadness from bodily cues and even more so when movement was involved as well. Dittmann et al. (1965) conducted studies in which they showed three series of short mute videos of an interview to participants to investigate the functions of bodily expression. One set of the videos depicted a female displaying a pleasant affect using both facial and bodily cues (Table 2.6), a second set showed the female depicting a homogeneous unpleasant affect, and the rest showed the female displaying pleasant facial cues but unpleasant bodily expressions.

Another factor introduced in Dittmann et al.'s (1965) study was through the method of presentation of the expressions. Some individuals were shown videos with the whole body and facial expressions while some were shown videos with masked faces. Individuals were then asked to rate the segments of videos on a pleasant-unpleasant continuum. Dittmann et al.'s (1965) results suggested that information

	Facial cues	**Bodily cues**
Pleasant	Smiles, laughs, direct gaze, raised head	relaxed posture with little movement
Unpleasant	Frowns and drawn tight expressions, lowered head	obvious muscle tension or fidgety nervous activity

Table 2.6: Cues used in Dittmann et al.'s (1965) study

about emotion is available in both facial and bodily cues but naturally cues were easier to interpret in the conditions where the individuals observed the body with the facial cues. This is in keeping with results reported by de Gelder (2006) where participants recognised the emotional state of an individual faster and more accurately, if the congruent behavioural cues were displayed. Interestingly, Dittmann et al. (1965) reported that even when facial cues were precluded, individuals were still surprisingly consistent in appraising feelings expressed in bodily cues.

There are three basic ways in which the body can be expressive. The first is through postural features. The second is through movement (and gestures). These two types of bodily expressions are generally termed *kinesics*. This includes all actions, automatic reflexes, posture and gestures, and other body movements. Gestures are intimately linked to speech and are important behavioural cues to emotional states and interpersonal relationships. However, gestures are not discussed within the scope of this thesis. The final type of expression in the body is the position in space of the body as a whole; where an individual stands relative to other people. This spatial use of the body is called *proxemics*. Spacing behaviour is unwittingly referred to frequently in daily communication for instance: '*I am close to X*' (Duck, 1998). The following sections discuss what can be expressed through the body and how.

2.4.3.1 Posture

Charlie Chaplin maintained that if an actor knew his emotion thoroughly he could show it in silhouette (Lasseter, 1987; Thomas and Johnson, 1981) while Hewes (1957) stated that "*... they (postures) speak an eloquent language in social intercourse. Most of us look to postural cues as well as facial expression and speech itself, in our never-ending efforts to interpret or evaluate people's motives, **moods** or behaviour*". Psychoanalysts have used a conscious knowledge of postural behaviour (while *lying* on a couch) to gain valuable insights into the psychodynamics of individuals. In the physical world, humans are capable of producing about a thousand stable postures which can be maintained comfortably for a period of time (Hewes, 1957). Even though no two body shifts are identical, it is possible to derive variants of postural behaviour that can be used interchangeably (Birdwhistell, 1971). The different parts of the body are attuned to each other, and the change of one partial posture leads to the rearrangement of the whole posture. Each deviant posture is meaningful and consists of a combination of different well-determined postures of single parts. The configuration of the resultant posture is characteristic to the individual and can also be regarded as an aspect of personality while the number and varieties of emerging deviations from the basic posture of an individual depends on the emotional situation at that time. Additionally Scheflen (1964) argues that an action, such as postural shifts leading to a movement away from others, is only significant in relation to the context it is being displayed in. Similar to results obtained later by Carroll and Russell (1996) on facial expressions, the potential emotive content of a posture varies considerably in accordance to the context in which it occurs or is perceived to have occurred.

Postural behaviour reflects an individual's emotional state with considerable agreement (Argyle, 1969; Dittmann et al., 1965; de Gelder, 2006) although results from Coulson's (2004) evaluative study indicated that some emotional states (anger, sadness and happiness) are easier to perceive through postures than others. De Silva et al. (2005) reported that participants from three different cultures (Japanese, Sri Lankan and American) agreed to a fairly large degree about the emotional state portrayed by postures but there were differences in how the intensity of emotion was perceived. This suggests that like facial expressions, some postural configurations may be universally recognised. However, like Coulson (2004), De Silva et al. (2005) reported that some features of postures (often emblematic) had specific meanings to specific cultures. Some cultures take great pains to ensure propriety in the manner of posturing in the right manner during public occasions maintaining careful distinctions on the basis of gender, age and social status (Hewes, 1957). Even in cultures where these mannerisms have been lost, individuals still unconsciously use postural displays to indicate various thoughts, attitudes and feelings.

Wallbott (1998) tested the quality vs. quantity theory on the bodily expression of emotions using 224 video recordings of actors and actresses portraying a wide range of emotions of varying intensity. The utterances used in the video clips were carefully designed to contain culturally neutral verbal content by creating two meaningless "*sentences*" out of typical syllabuses from six languages: "*Fee gott laish jonkill gosterr*" and "*Hat sundig pron you venzy*". The video clips were then *coded* to record the body movements and postures displayed by the actors in the clips using a set coding schemata. Wallbott (1998) was able to single out distinct behavioural cues specific to certain emotional states. For instance, an erect posture was rare in cases of portraying shame, sadness or boredom rather the actors used a collapsed body posture. A posture with raised shoulders was typical in cases of anger but moving the shoulder forward was seem to portray fear and disgust. The most significant differences were found in hand and arm movements. For instance lateral movements of the arm and arms stretched out in front were associated with anger, an *active* emotion; crossing the arms in front were associated with pride and disgust; opening and closing of the hands was associated with anger and fear. In keeping with earlier studies (James, 1932; Mignault and Chaudhuri, 2003; Schouwstra and Hoogstraten, 1995), the position of the head differed significantly between emotions (Wallbott, 1998).

It seems that the postures of some parts of the body are more important than other. For instance, James (1932), found that a forward head and trunk position is taken to mean an emotional expression whereas the outward stretch of the arms and hands suggest a movement and not an emotion. In addition, one part of a total postural configuration is noted at the expense of the remaining posture. For instance closed/clenched fists indicate tension and the expression of anger. The head and the trunk of the body were found to be the most significant for the generic expression of the total posture. Secondary factors which were found to be important included the expansion/contraction of the chest and the position of the shoulders (raised/drooped) (James, 1932).

Until recently, although the role of body posture in the communication of emotional states had been discussed theoretically, little empirical study regarding the attributes of posture and the computational modelling of these attributes in virtual humans have been conducted. In comparison to facial expressions and gaze behaviour, literature on defining cues to the bodily expression of emotions is incredibly scarce and mostly made of vague descriptions. This presents problems when trying to computationally model the display of affect through nonverbal behaviour in full-body virtual humans. Darwin (1872) presented some observations on the bodily expression of emotion and these have often been cited in various other studies including the work conducted by Wallbott (1998) and Coulson (2004). Table 2.8 gives cues based on Wallbott's (1998) post-hoc comparison with Darwin's (1872) original observations summarised in Table 2.7. Most empirical research tend to involve recordings of body movement which confound findings since the effects could not be attributed to the posture alone.

Emotion	**Darwin's (1872) observable bodily cues**
Anger	Whole body trembles, clear intent to push or strike, gestures become purposeless and frantic, pacing, shaking fist, erect head, well expanded chest, feet firmly planted on the ground, elbow(s) squared or arms rigidly suspended by the sides, clenched fists, squared shoulders
Sadness	Motionless, passive, head hangs on contracted chest

Table 2.7: Behavioural cues to some emotions as cited in Wallbott (1998)

Emotion	**Wallbott's coded bodily cues**
Anger	Erect upper body, raised shoulders, forward lean, arms stretched out front, lateral movements of the hands, pointing gestures, opening/closing of hands, high movement activity, expansive movements, high movement dynamics
Sadness, Despair	Collapsed upper body posture, forward shoulder positions, opening/closing of hands, low movement activity, expansive movements, low movement dynamics

Table 2.8: Behavioural cues to some expressions as cited by Wallbott (1998)

Coulson (2004) conducted an evaluative study on static non-emblematic postures in order to ascertain the importance of anatomical attributes in conveying an emotional state. Coulson (2004) showed participants a stimuli set chosen from 176 postures displayed on simple computer generated stick figures. The stimuli set was limited to the Ekman's (1982) six basic emotions: *anger*, *disgust*, *fear*, *happiness*, *sadness* and *surprise*. Each image of the stick figure was generated from three angles to determine if the recognition rates differed between viewpoint. Participants were presented with a stick figure displaying each static posture on a desktop PC and asked to select which of the six emotion labels best suited the posture. In general, Coulson (2004) found that the postures were easier to recognise from the front. Coulson (2004) also reported that anger and sadness along with happiness were attributed to large numbers of postures with agreement rates comparable to those obtained in similar studies on facial expressions. However, since the only positive emotion in the label set was happiness, participants might have attributed non-negative postures to happiness. The postures chosen by participants to best portray anger and sadness were in keeping with the descriptions reported by Darwin (1872).

Kleinsmith et al. (2005) conducted a statistical analysis of emotional posture produced by Japanese actors. Using Multidimensional Scaling they found three main dimensions that explained the variation of posture. They interpreted the first as corresponding to *arousal* that separated sadness, depressed and upset (low arousal) from fear, happiness, joy and surprise (high arousal). Low arousal postures tend to have a bent head and arms placed to the side of the body. The second dimension corresponded to *valence* and separated surprised and fear (low valence) from happiness and joy (high valence). Low valence postures consisted of the head bent forward and the hands raised in front of the face. High valence postures had a raised head and hands held high and away from the body. The final dimension seemed to represent an *action tendency*, with anger being an active emotion while fear and surprise were passive (low action tendency). In passive postures, the hands were raised near the face and the body was kept narrow; whereas in active posture the elbows were bent out to the side of the body and the hands were kept around the hips. While it could be argued that the dimensions Kleinsmith et al. (2005) found do not correspond exactly to arousal and valence, the model is consistent with other findings. In other studies a lowered head and bent forward trunk was found to correspond to submissiveness and negative emotions such as sadness, shame and humiliation (Coulson, 2004; Darwin, 1872; Wallbott, 1998; Mignault and Chaudhuri, 2003; Schouwstra and Hoogstraten, 1995). On the other hand an upright posture and raised head indicated dominance and positive emotions such as pride and joy (Coulson, 2004; Darwin, 1872; Wallbott, 1998; Mignault and Chaudhuri, 2003; Schouwstra and Hoogstraten, 1995). In studies associated with Kleinsmith et al. (2005), De Silva et al. (2005) found that posture could also be used to distinguish between different nuances of similar emotions, for instance, the differences between *joy* and *happiness* were particularly notable and consisted mostly in the openness of the arms and distance between the feet. There are exceptions, for instance, Mignault and Chaudhuri (2003) mention that in some cases, a lowered head indicated an *illusionary* smile which was interpreted as submission or joy. This shows the complexities involved in modelling nonverbal behaviour. Kleinsmith et al.'s (2005) model is very recent and yet to have an impact on virtual humans, however, the postures described by Coulson (2004) were used in an experiment discussed in Chapter 5.

In addition to portraying emotion, postural behaviour plays an important role in conversation management and representing relationships (Argyle, 1998). For instance, individuals adopt postures in accordance to their affiliation for the other (Argyle, 1969). If an individual dislikes the other, they typically stand *hand on hips*. Scheflen (1964) argued that configurations of posture or body positions indicate the dynamics of an interaction, are used unconsciously, and are reliable indicators of certain aspects of interpersonal communication. For instance, postural shifts leading to a movement away from others is often seen to indicate completion and temporary disengagement from the interaction. High affiliation (liking) is expressed through postures that bring people closer (e.g. leaning forward while sitting), while low affiliation is expressed by closed posture that present a barrier between people (e.g. crossing arms). Bécheiraz and Thalmann (1996)'s agents were able to choose different postures that displayed different levels of affiliation. Another dimension of attitude is status: *dominance* and *submission*. Dominant people tend to have large open postures, that appear to increase their size, such as standing straight, expanding the chest and putting hands on hips, while submissive people tend to have small, closed and hunched over postures. Individuals of higher status sit in front of individuals while those of lower status adopt a less open and direct sitting posture (Argyle, 1969). Gillies and Ballin (2003) used a model of both affiliation and status to model social interactions between virtual humans, however, their model is still to be evaluated in a multi-party conversation.

Scheflen (1964) identifies three dimensions of posture during social interaction: spacing, orientation and positioning.

Non-inclusiveness - Inclusiveness: Individuals in an interaction tend to define group space by the placement of their bodies. If a dyad is not private, a larger interpersonal distance is maintained and the individuals stand outwards at either a 60^o or 90^o angle (Scheflen and Scheflen, 1972). If the dyad is private, the individuals stand closer and face each other sometimes using their arms as barriers.

Vis-à-vis - Parallel: In an interaction, individuals can either situate themselves face to face (*vis-à-vis*) or side by side (*parallelism*) (Scheflen and Scheflen, 1972). Individuals situate themselves vis-à-vis usually to in an interaction thought of as involving an exchange of information: teaching, informing, conversing, and quarrelling. In contrast, parallelism is used when individuals are involved in an interaction toward some third party: two individuals quarrelling against a third and sharing in reading.

Non-congruence - Congruence: The ways between which the bodies of individuals in an interaction are arranged complimentary to each other is termed *postural congruence*. Postural congruence indicates similarity in views, or roles in the group and gives an indication of status (Scheflen, 1964). Congruence can occur in two ways: direct or mirrored. In direct congruence, individual hold the exact same posture while in mirrored congruence, individuals hold a mirror-image of each others' posture.

In addition to the shape of the posture itself, the timing of shifts between postures is another manner in which the individual's internal state is displayed. Egges and Magnenat-Thalmann (2005) recorded 10 participants in conversation as a means of obtaining motion data. In the recorded data, they reported that the three most common types of idle behaviour were posture shifts, continuous but smaller postural variations due to breathing or maintaining balance, and supplemental idle motions such as touching of the face (Egges and Magnenat-Thalmann, 2005). The first type, posture shifts, is closely linked to *interactional congruence* (Kendon, 1970). During a conversation, individuals' posture shifts tend to become synchronised with each other, and with the rhythms of each others' speech thus increasing perceived sense of rapport. Female individuals display more direct body orientation (vis-à-vis) and gaze with each other while male individuals engage in forward leaning and postural congruence with increasing affiliation (Duck, 1998). Postural congruence occurs more in individuals of the same gender than in opposite-sex dyads (Duck, 1998). Postural congruence is such a common occurrence in social interactions that individuals with long-term ties often shift postures in congruence when they are temporarily arguing to communicate a sense of their continuous friendship (Scheflen and Scheflen, 1972). In a vis-à-vis oriented interaction, a lack of congruence is noticed when the individuals are not of equal stature. Scheflen's (1964) three dimensions of postural relation between individuals occur simultaneously. The synchronisation of posture is also likely to be due to a deeper synchronisation of each individual's posture with the rhythm of speech. Cassell et al. (2001) have studied how posture shifts relate to discourse structure. They found that posture shifts tend to happen most at the start of discourse segments (change in topic of conversation) and during turn taking. They used these results to implement the behaviour in an agent - Rea (Figure 2.1(d)). These aspects of postural behaviour can be simulated in virtual humans by using a head tracker to detect posture shifts and a microphone to detect the start and end of speech in the interactant (Gillies and Slater, 2005).

Animating posture is relatively straightforward if subsequent postures are not widely varied. It is mostly done using a library of pre-defined poses from which an appropriate posture is chosen based on some of the factors listed above. An example of such a system is presented by Guye-Vuillème et al. (1999). Transitioning between postures can use a standard motion transitioning method. In order to make sure there is enough variety in posture some randomness is used when choosing postures. For even more variety Gillies and Ballin (2004) choose multiple postures at a time from the library and interpolate them to generate new postures, using random weights (Figure 2.1(g) and Section 2.4.3.2 for a discussion of motion interpolation). However, it is important to ensure that the multiple postures chosen have a significant number of common attributes in order to avoid creating a resultant posture of a completely different meaning. The angry and sad postures described by Coulson (2004) were interpolated using an algorithm similar to Johnson (2003) and then used extensively in one of the experiments reported in this thesis. The experiment was designed to investigate the impact of affective agents posed vis-à-vis on participant responses (Chapter 5 and Figure 5.9).

2.4.3.2 Quality of movement

Many researchers believe that bodily movement is a highly accurate cue to the emotional state of an individual even in the form of dynamic point light displays (Johansson, 1973). Wallbott (1998) argue that there is a distinct pattern in the postural behaviour associated with at least some emotions both in the qualitative and quantitative sense. Another important factor is the level of movement activity associated with the emotions. For instance, energised movements were typical of joy, anger and fear in that order while less movement was associated with despair and shame (Wallbott, 1998).

Wallbott's (1998) studies were focused on exploring behavioural cues by asking actors to act emotional expressions out and then coding out the behaviours. On the other hand, like Dittmann et al. (1965), Montepare et al. (1999) investigated the extent to which individuals perceive bodily expressions of emotions. Montepare et al. (1999) showed participants a set of three seconds muted video dramatisations of two actors in behavioural scenes depicting one of four emotional states: *anger*, *sadness*, *happy* and *neutral*. Two actors were used in each scene in order to increase the naturalness of the dramatisations, however, this could have had the effect of confounding the results reported since it provided participants with more knowledge about the context (Argyle and Trower, 1980; Ekman, 1965; Scheflen, 1964). In addition the actors were given the freedom of a loose script and the space to improvise, however, the focus of the study was on the bodily cues provided by the movements of the actions. This helped in making the portrayal by the actors more natural. The faces and dialogues of the actors in the clips were blurred and muted before being used in the evaluative part of the study. Participants were asked to record the dominant emotion perceived in the clips and rate the clips with respect to characteristics of body movement on a set of six 7-point response scales: *smooth-jerky*, *stiff-loose*, *soft-hard*, *slow-fast*, *expanded-contracted*, and *no action-lot of action* (Montepare et al., 1999). Results indicated that neutral clips were identified with a higher degree of accuracy than emotional clips. Amongst the emotional clips, angry clips were identified more accurately than sad or happy clips both of which were identified with similar levels of accuracy. Figure 2.2 depicts the results obtained from the study on the characterisation of body movements with respect to the emotion categories used. Angry clips were characterised by individuals to be jerky, stiff, fast, hard, expanded and full of actions. In addition, angry displays of emotion were recognised with the most accuracy (Montepare et al., 1999). This is in agreement with the *face-in-the-crowd* effect discussed in Section 2.5.2.1. This is also in keeping with results reported later by Coulson (2004).

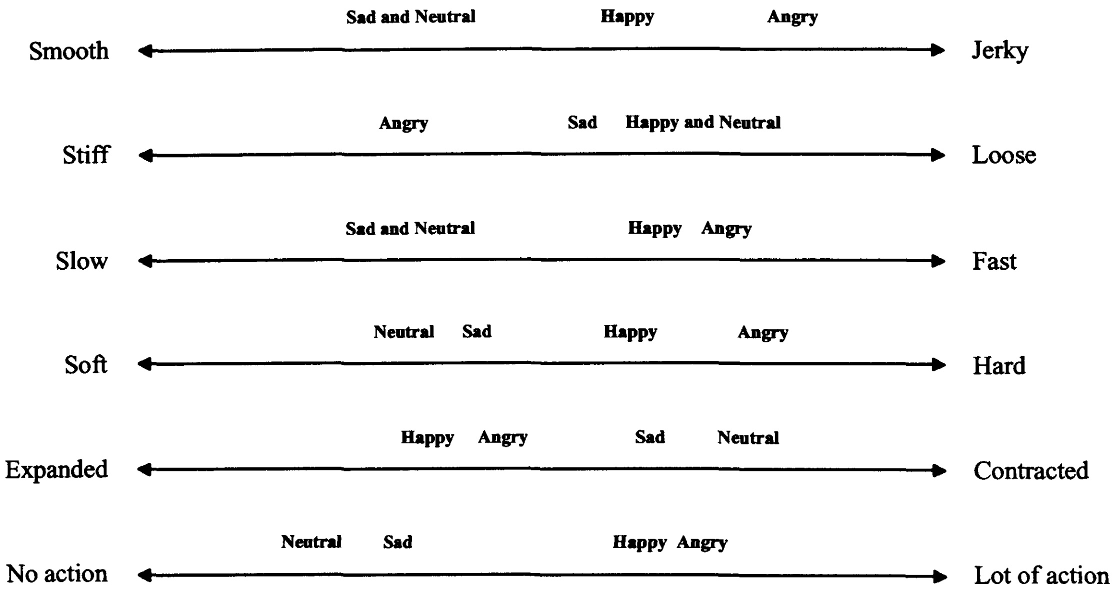


Figure 2.2: Montepare et al.'s (1999) ratings of emotional cues with respect to body movement

Elsewhere Paterson et al. (2001) reported studies in which participants were shown arm movements of activities such as eating, drinking, lifting and knocking, obtained from actors posing the movements in one of 10 internal states including *anger*, *sadness* and *neutral*. Analysis of the movements suggested a correlation between the emotion categories and the velocity of movements (Paterson et al., 2001). In keeping with Montepare et al.'s (1999) results, sad movements were always slower than neutral while angry movements were faster than neutral. Energetic movements were positively correlated with shorter durations and greater magnitudes of average velocity, peak velocity, acceleration, deceleration, and jerk. Angry movements had the shortest duration and highest velocities while sad movements have the longest duration and lowest velocities. Paterson et al. (2001) carried out a categorisation study with participants who viewed the movements in form of point light stimuli. Participants were able to correctly recognise angry and sad movements even with limited cues, supporting the premise that body movement plays a key role in the communication of affect. When the original movements were modified temporally, angry movements of various speeds were categorised as having differing intensities of anger while sad movements were categorised as angry movements when sped up. This suggests that individuals are especially heightened to recognise specific attributes in the angry actions even if they are modified temporally. In addition, the speed of body movement seems to be key to creating a sad motion. This result is in keeping with the earlier studies conducted by Montepare et al. (1999) and Coulson (2004) in which participants recognised anger from posed bodily expressions with high accuracy.

Animating expressive body movement involves taking a particular type of motion such as walking, sitting down or drinking, and applying a style to it. A style can represent many things. It can be an emotional style of movement, or an individual's personal style of movement, which might in some way reflect their personality. Therefore the same animation techniques can be applied to many of the different factors discussed. Bruderlin and Williams (1995) do a frequency analysis on pieces of motion and are able to alter the relative importance of various frequency bands. However, these transforms do not capture the full subtlety of emotional movement. Chi et al.'s (2000) EMOTE models motion transformations based on Laban Movement Analysis (LMA) which is a method of studying, observing, describing, notating and interpreting "human movement"(Davies, 2001). They implemented two components of the system: *Shape* (the changing form the body makes in space) and *Effort* (how the body concentrates its exertion while performing an action). The shape component alters the amplitude of the motion in the horizontal, vertical and side-to-side (sagittal) planes relative to the body. The Effort transform is more complex but involves altering the continuity and timing parameters of the motion paths and adding flourishes which add to the expression such as wrist bends. However, even though the system allows for the creation of natural movements for virtual humans, the Shape and Effort components are not linked to an easily comprehensible interface which allows for the generation of an emotive action.

Another approach is to *learn* a style from one or more examples of a style and then applying it to a new, neutral motion. This was done successfully by Paterson et al. (2001) in point light stimuli sets of body movement. This relates to the non-parametric method for creating gaze behaviour discussed in Section 2.4.2. The data typically comes from the captured motion of an actor performing an action in a given style. A variety of different methods have been used to do this learning. One of the first attempts at this was presented by Amaya et al. (1996). They learnt two parameters that represented a style: a *temporal* transform and a *spatial amplitude* transform. More recently, researchers have tried a number of different representations of style such as Brand and Hertzmann's (2000) Hidden Markov Models and Hsu et al.'s (2005)'s linear matrix methods. Liu et al. (2005) present a more sophisticated model that uses a physically-based simulation of human movement with a number of parameters that represent style. They use an optimisation method to learn the parameters from input data.

On the other hand, Rose et al. (1998) do not attempt to extract a style that is independent of the type of action but still use a set of sample data to build a library of different styles for a single action. They define a *verb* which consists of a set of sample motions representing a single type of action in different styles. They then define a number of *adverbs* which are numerical descriptors of a style of motion, for instance, sample x might be 90% *happy* and 20% *surprised.* Given a new set of adverb values they can generate a new motion by interpolating the example motions using radial basis functions. Mukai and

Kuriyama (2005) suggest using geo-statistical interpolation to combine motions as they are better able to represent the statistical properties of motion. This type of interpolation method requires some way of capturing large amount of motion data with associated adverb values, such as Garcia-Rojas et al.'s (2005) proposed motion capture and annotation system.

Again, the main advantage of using parametric methods of modelling body movement is the compactness of the result. However, body movement introduces a much larger degrees of freedom caused not only by the three-dimensional aspect of skeletal animation but also due to the variety of context-dependent postures and gestures involved in portraying affect. The major disadvantage of all these methods is that it is not clear that the style of motion can be completely separated from the action being performed. It is not clear that what makes an angry walk angry is the same thing that make an angry gesture angry. The most significant challenge of building a computational model of kinesics is the lack of information describing what aspects of a motion/action portrays a specific emotion. This thesis focuses on building a parametric model of kinesics (Chapters 5 and 6) which can be used to portray two emotional states: *Anger* and *Sadness*.

2.4.3.3 Proxemics

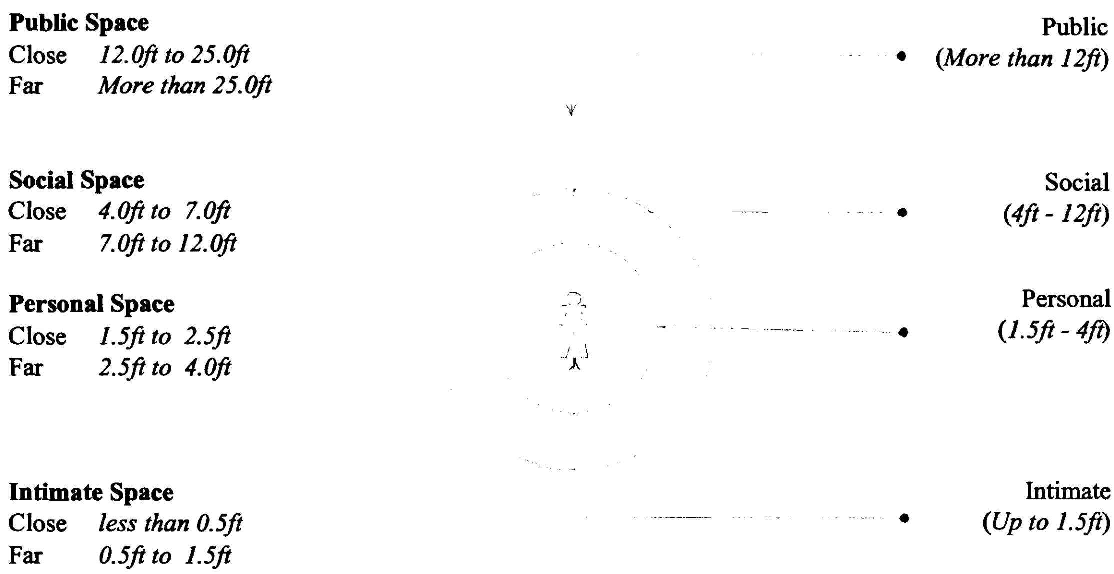


Figure 2.3: Hall's (1969) concentric (not necessarily circular) zones. 1 foot $\approx$ 0.305 metre

The flow and changes in the interpersonal distances between individuals in a conversation is an integral part of communication (Proxemics). Hall (1969) presented a theoretical model of proxemics that consists of four concentric zones. Each concentric zone commands a different level of social intimacy which controls the interpersonal distances maintained by the individual with others (Figure 2.3). Impersonal business and social gatherings normally take place in the close phase of the social distance zone. Further information regarding Hall's (1969) model is available in Section I.2 of Appendix I. Hall's distances are averages for middle class American adults. Culture has a strong effect on spacing behaviour. For instance, English males keep further apart that French or South American males (Argyle and Trower, 1980; Scheflen and Scheflen, 1972; Sommer, 1969). Spacing behaviour also varies with personality and environmental factors, context and emotional state. For instance, an angry person who is emphatic about making a certain point will move in closer to the other and turn up the vocal volume (Duck, 1998).

Spacing behaviour is also interconnected with other modalities of expression such as gaze behaviour. The appropriate interpersonal distances maintained between individuals is held steady by balancing increasing proximity with reduced eye contact in keeping with Argyle and Dean's (1965) *theory on equilibrium*. The lower limit of the distance is determined by physical contact whereas the upper limit is defined by factors of visibility and audibility (Argyle, 1969; Hall, 1969). Equilibrium is only held within the personal distances of the individual. If individuals see each other from further away, they move towards each other but break eye contact until they are closer (Kendon, 1967). However, this behaviour is dependent on the interpersonal relationship between the two individuals. Two strangers approaching each other will look towards each other at a distance of 12.0ft or more in acknowledgement of their presence and to avoid a collision but will avert their gaze at closer distance (polite inattention) (Scheflen and Scheflen, 1972). Goffman (1963) termed this behaviour: *civil inattention*. Argyle and Dean (1965) conducted a set of studies with dyads separated at varying interpersonal distances of 2.0ft, 6.0ft and 10.0ft. Dyads in the 2.0ft conditions leant backwards while conversing and those in the 10.0ft conditions leant forwards.

The equilibrium theory was tested by Bailenson et al. (2001) in an IVE study based on two factors: level of visual realism of the agent and level of mutual gaze engaged. Participants were required to walk towards and around a *male* agent in a virtual room. Participants maintained more space around the agent in comparison to other virtual objects. Female participants maintained more interpersonal distances with the agent in the condition where the agent engaged in mutual gaze (Bailenson et al., 2001). Male participants showed no distinction. Bailenson et al. (2003) later extended this work to include head gestures and female agents. In addition, individuals in one condition were introduced to a confederate and made to believe that the agent was in fact an avatar of the confederate. This was to introduce the factor of perceived agency and investigate if participant's response to the agent differed if the participant believed that the agent was representing a real person. There were two interaction effects: agency with gaze and agency with participant gender. The results from the original experiment (Bailenson et al., 2001) were replicated in that participants maintained greater interpersonal distance with the agent that engaged in mutual gaze. This difference did not show in the conditions in which the individuals thought of the agent as an avatar (Bailenson et al., 2003). In other words, higher behavioural realism was not necessary for personal spacing behaviour to be maintained in the case of perceived avatars where participant thought that the agent was representing the confederate. This in contradiction to results obtained in Nowak and Biocca's (2003) study which indicated that individuals responded socially to the virtual human regardless of agency in the same manner. This might be due to the nature of interaction in Nowak and Biocca's (2003) study where the participant exchanged information with the virtual human in a very scripted and artificial manner[6]. Secondly, female individuals maintained greater

[6]See section 2.1 for more details on Nowak and Biocca's (2003) study.

interpersonal distances from the perceived avatar than the agent (Bailenson et al., 2003). Males did not show this distinction. In addition the size and shape of the personal space bubble (frontal distance larger than back) maintained around the virtual human was similar to that maintained in the physical world (Bailenson et al., 2003).

Elsewhere, Jeffrey (1998) observed in an ethnomethodological type study that the avatars of participants in an online VE maintained a distinct physical distance between each other in a face-to-face conversation which correlated with group size. The study also uncovered that spatial invasions produced anxiety arousing behaviour (Jeffrey, 1998) with attempts at re-establishing equilibrium in keeping with Argyle and Dean's (1965) theory. Bailenson et al. (2003) have used the results of these studies to argue for the utilisation of behaviours, especially spatial behaviour, as a realism gauge in IVEs. Not only is spacing behaviour an integral part of social interaction, it is easily recordable in VEs making it a potential tool to measure participant responses to virtual stimuli. Despite this potential, there has been little in the way of work on virtual humans that display proxemic behaviour. Gillies and Slater's (2005) agent maintains a realistic social distance, but this work is not integrated with their earlier work on attitude (Gillies and Ballin, 2003). More recently, Bickmore and Picard (2006) presented a clever implementation of proxemics in a desktop based system. They use close ups of an agent to indicate close proximity when discussing intimate subjects.

2.5 Responses to expressions of emotion

The previous sections dealt with computational models that could be used to generate emotional states in virtual humans and the different behavioural cues that can be used by the virtual human to display these states. The following sections discuss the different responses to expressions of emotional states as observed in the physical world. This section aims to review the different types of responses that could be expected when participants interact with affective virtual humans.

The determinant of an individual's response to a given situation is dependent on a number of factors including their cognitive appraisal of the situation and it's social context (Schachter and Singer, 1962). In this section a brief review of possible responses to emotional situations is reported with particular focus on one *basic* emotion: Anger. The purpose of using this particular emotional state in this thesis is three-fold. Anger suffers relatively less controversy in the literature with respect to categorisation and its status as an emotional state. Anger has been associated with a specific physiological response pattern (Section 2.5.2.3) and is therefore regarded as a bona fide emotion in many of the models discussed in Section 2.3.5. The second reason is due to its association with threat perception. There is evidence that individuals are particularly sensitive and respond to others displaying cues that portray anger (de Gelder, 2006; Green et al., 2003). Finally, anger is the most accurately recognised emotional state (Bartneck and Reichenbach, 2005; Coulson, 2004; Paterson et al., 2001). This is highly desirable in experimental scenarios since it favours the successful recreation of an emotive virtual human by incorporating minimal behavioural cues.

2.5.1 Emotional contagion and feedback

Moods and emotions of others around an individual influence their emotional state and subsequent behaviour. There is also evidence that behavioural feedback influences the emotional state of individuals and postural feedback may intensify the emotional experience (Duclos et al., 1989; Flack Jr et al., 1999; Hatfield et al., 1994). This process of *catching* the prevailing mood of others is known as "emotional contagion" (Doherty, 1997). Schachter (1959) argued that physiological arousal evoked by a threat causes uncertainty, and therefore individuals in similar threatening situations desire to affiliate themselves with others to evaluate the intensity, nature, and appropriateness of their emotional state. In an empirical study involving the virtual recreation of a bully-victim type situation, young participants felt empathy towards the victim and felt anger towards the bully especially when the participant identified with the victim (Paiva et al., 2004, 2005; Hall et al., 2005). Hall et al. (2005) reported that expressions of empathy were increased when young participants had high levels of interest in the conversation between the agents. In other words, the more young participants perceived that a conversation was believable, the more they felt sorry for the agent. Further results indicated that if the young participants perceived that they were similar to the agents (identified with them), they expressed greater empathy and liking (Hall and Woods, 2005). The concept of the emotional similarity hypothesis might prove to be especially relevant under collaborative scenarios of a particularly stressful nature such as those explored by Gratch and Marsella (2004) or in situations calling for strategic collaboration in games. This property can also be used as a possible participant response to an angry virtual human. For instance, if a participant observes an agent displaying appropriate behavioural cues against another agent (victim) in an IVE, will the participant develop feelings of sympathy towards the victimised agent?

2.5.2 On responses to Anger

Anger could be the emotional response to events perceived as an injustice, a threat, frustrating and a struggle for dominance amongst others. Anger is displayed by individuals to curtail the behaviours of others in situations where they have broken the social rules or expectations. Anger is not unlike any other emotion in that its expression and the response it invokes differs from one individual to the next. Perceived anger in others encourages individuals to behave in a manner that either curtails the situation or exacerbates it depending on the context, personality and emotional state of the individual (Brewer and Hewstone, 2004). The following sections discuss well-known responses to expressions of anger in others.

2.5.2.1 Subconscious response: Focusing attention

There is evidence that threat-related facial expressions such as anger, invokes individuals to use different viewing strategies. Green et al. (2003) conducted a study to investigate the cognitive processing of different facial expressions. Results indicated that the visual scanpaths of individuals who looked at threat-related faces (angry/fearful) were characterised by increased distances between fixations (*extended scanning*), with longer and exaggerated focusing on certain features areas of the face (eyes, nose, and mouth) especially in the case of angry facial expressions. Green et al. (2003) speculated that the longer saccadic eye movement might reflect heightened autonomic responses to threat-related stimuli while the increase in focus might facilitate cognitive appraisal of the possible ramifications of the stimuli to the individual. In a series of earlier studies, Hansen and Hansen (1988) reported that threatening faces *pop out* in a crowd. In a photograph of a neutral or happy crowd, angry faces were found more efficiently, with faster reaction times and lower errors in comparison with neutral or happy faces in an angry crowd. However, Fox et al. (2000) criticised Hansen and Hansen's (1988) studies by arguing that some of the photos with an angry face had notable skin markings which might have confounded the results.

Fox et al. (2000) recently revisited this issue to investigate the *face-in-the-crowd* effect using simple drawings of faces (schematic faces) and concluded that the detection of an angry face is fast and efficient. Fox et al. (2000) carved out two hypothesis for their study in accordance to two theoretical premises: the *emotionality* hypothesis and the *threat* hypothesis. If individuals were given two sets of faces, all angry or all happy, in accordance to the threat hypothesis, responses to a task concerning the angry faces would be slower than the happy faces since angry faces demand a longer *dwell* time than happy faces. The emotionality hypothesis would not differentiate between the sets. In addition, if there were a discrepancy in the set of faces where by a different facial expression was placed in the midst of an otherwise uniform neutral set of faces, the threat hypothesis states that an angry face would be detected faster and with fewer errors than a happy face. The emotionality hypothesis would once again make no distinction. Studies with varied exposure times, indicated that the detection of anger held more priority than the detection of happiness thereby supporting the threat hypothesis. Fox et al. (2000) also confirmed that the detection of an angry expression is indeed faster and more efficient. Further studies with inverted facial expressions confirmed that the effects observed in the prior studies were indeed due to the angry expressions.

Mogg and Bradley (1999) presented a study in which participants were shown pairs of the same face (one neutral and the other either happy or angry) side-by-side for a short while. The faces were then replaced with a set of neutral *masked* faces. Participants were asked to press buttons to indicate the neutral masked face in which a probe was displayed. Individuals were faster to react to the probe when it was on the side where the threatening face was previously displayed in comparison to the happy face. This suggested that unconsciously individuals focused attention on threatening cues. These results are reminiscent of the results reported by Montepare et al. (1999) in which participants recognised bodily displays of anger more quickly and accurately than happy or sad (Section 2.4.3.2). The *face-in-the-crowd* effect seems to be in keeping with the Darwinian theory which argues that emotions are a remnant of the evolutionary process to aid the fight or flight phenomena in the face of a threat (Darwin, 1872). This tendency to focus attention towards more threatening facial expressions is greater in high trait anxious individuals (Mogg and Bradley, 1999).

These studies indicate that there is a heightened processing of threatening faces, a deployment of selective visual attention and maybe even heightened sensitivity to detect variations in the intensity of an angry expression. Therefore, participants in IVEs are likely to accurately recognise the emotional state of a virtual human displaying behavioural cues associated with Anger.

2.5.2.2 Psychological responses: Stress and Anxiety

Stress and anxiety are two related but distinct forms of responses to a threatening situation. Stress is an individual's response to any situation or event that requires adjustment or change (Lazarus, 1993). It is therefore the natural response to any situation perceived to be a threat to an individual. Anxiety is the multi-system response to a perceived threat. It reflects a combination of biochemical changes in the body, the individual's personal history, memory, and the social context. Though anxiety is related to fear, it is often unfocused, vague, and hard to pin down to a specific cause. Traditionally anxiety is considered as a set of two psychological variables: *State* anxiety (S-Anxiety) and *Trait* anxiety (T-Anxiety). S-anxiety is defined as an unpleasant emotional arousal in the face of threatening stimuli while T-anxiety reflects the differences in the tendency to respond due to individual traits. The level of state anxiety generated in an individual to any give situation depends on their general level of trait anxiety. In other words, the same stimuli might be viewed as anxiety-inducing and perceived as a threat on differing levels by different individuals.

Feldman et al. (2004) report a study in which preparing to speak in public (anticipatory stress) led to greater threat appraisal, a negative emotional state and increased physiological arousal than preparing to read out aloud. This effect was especially noticeable in speech anxious individuals. Anxiety also impacts on the ability to identify the facial expressions of specific emotions. For instance, Gard et al. (1982) reported that participants with low trait anxiety were more accurate in interpreting facial expressions of emotions from a set of slides under stressful conditions where state anxiety kicks in. However, participants with high trait anxiety were more accurate in interpreting the emotional expressions under non-stressful conditions. Individuals appraise the extent to which stimuli in their environment are threatening (primary appraisal) given their resources for coping (secondary appraisal). If the extent of the perceived threat exceeds their resources for coping, individuals react with greater negative emotion associated with threat perception (Feldman et al., 2004). This has been explored by Gratch and Marsella (2005) in order to create realistic behaviour in virtual humans under stressful scenarios.

Mullins and Duke (2004) focused on response times to identifying facial expressions and concluded that in non-threatening situations, participants with high trait anxiety are slower to identify angry emotions. They argued that this might be due to the similarities in facial cues between emotions at low intensity in keeping with the results reported by Hess et al. (1997) and Bartneck and Reichenbach (2005) as discussed earlier in Section 2.3.3. However, Mullins and Duke (2004) also reported that socially anxious participants were faster at detecting threat-related (angry and fearful) expressions while experiencing moderate levels of anxiety when compared to a very high or a very low level of anxiety. They speculate that at intermediate levels of arousal, experiencing state anxiety produces a hyper-vigilance that aids individuals in identifying threatening expressions. However, as levels of anxiety increased further, the cognitive symptoms of social anxiety interfered with response times (Mullins and Duke, 2004). In any case, it is generally accepted that a cognitive appraisal of threat is a prerequisite for experiencing anxiety (Lazarus, 1991).

2.5.2.3 Physiological responses: Arousal

There is evidence that specific emotional states may be associated with characteristic physiological responses (Ekman et al., 1983; Vyzas and Picard, 1999). For instance, there are detectable heart rate increases during excitement, mental concentration and during the presentation of intense sensory stimuli (Frijda, 1986). Due to the discrepancies in defining and categorising emotions (Kleinginna and Kleinginna, 1981) and ongoing work in physiology (Scheirer et al., 2002), building an exact mapping is still work in progress (Lisetti and Nasoz, 2002). Researchers have studied the relationship between psychophysiology and emotions under four possible categories: *one-to-one*, *one-to-many*, *many-to-one* and *many-to-many*. Scheirer et al. (2002) argue that physiological measurements alone are not adequate to provide a coherent insight into the individual's emotional state and suggest viewing results in tandem with more than one physiological measure and a behavioural measure. On one hand, there are those who believe that each emotion has a unique autonomic signature and on the other, those who believe that all emotions are accompanied by the same state of nonspecific autonomic arousal which only varies in intensity. However, it is the general consensus that physiological signals such as galvanic skin response (GSR), heart rate (HR), respiration and muscle tension provide valuable insights regarding the intensity and quality of an individual's internal state (Collet et al., 1997).

Luborsky et al. (1963) reported that participants who look at threatening pictures respond differently to participants viewing pictures of neutral content. Their studies also indicated that GSR responsiveness is positively correlated to avoidant patterns of eye fixations. Participants looking at the threatening pictures had high peaks in their GSR responses associated with shorter eye fixations (Luborsky et al., 1963). Participants also looked longer at the less threatening background of the pictures. Ekman et al. (1983) argued that although GSR alone failed to distinguish between all the Ekman's (1982) basic emotions, autonomic responses can distinguish not only between positive and negative responses but also between two specific negative emotions: anger and fear. Changes in HR associated with anger, fear and sadness were all significantly higher than those for happiness, surprise and disgust (Ekman et al., 1983; Levenson et al., 1990). Heart rate is reported to increase most during fear, anger, sadness, happiness, surprise and finally disgust in that order. More recently, Brosschot and Thayer (2003) concluded that heart rate activation associated with a negative emotional episode was of longer duration than HR activity related to a positive emotional activity. This stage of prolonged increased HR and low heart rate variability (HRV) is known as a "*hyperkinetic state*" (hypertension). The effect was due to emotional valence (emotion negativity) not arousal (intensity). Brosschot and Thayer (2003) argue that a longer HR response might be due to slow recovery after stress. Collet et al. (1997) extended this research by conducting another study with the objective of investigating the specific physiological responses to emotional stimuli. The physiological responses studied included skin conductance, respiration and skin blood flow as well as skin potential, skin resistance and skin temperature. The stimuli were categorised in accordance to Ekman's (1982) basic emotions. Collet et al. (1997) concluded that a group of three physiological signals (skin conductance, thermo-vascular and respiratory responses) taken as a whole were sufficient to separate each emotion.

Studies into the physiological response to emotional stimuli is still in it's exploratory stages, however, they can provide a valuable and objective method of exploring participant responses to expressive virtual humans. This methodology has already been used successfully by Meehan (2001) to explore participant responses to stressful VEs (a virtual precipice). On a more subtle level, Prendinger et al. (2005) conducted a study to investigate the effect of expressive virtual humans on the affective state of participants as measured through questionnaires and physiological responses. Participants were asked to take part in a quiz in which an agent (Shima) gave instructions. As part of the experimental design, a delay in giving some of the questions in the quiz was deliberately programmed in order to frustrate the participant. The participants were randomly assigned to one of two conditions: *affective* and *non-affective* (control). In the affective condition, the agent responded with 'happy-for' or 'sad-for' expressions depending on whether the participant gave a correct or incorrect answer. Also in the affective condition, the agent apologised, verbally and nonverbally, after the deliberate delay. In the non-affective condition, only verbal indications of the right or wrong answer was given. In addition, the agent did not offer an apology for the deliberate delay. Prendinger et al. (2005) reported that an empathetic agent had a positive effect on the participant's perception of the difficulty of the task and significantly decreased stress caused through the delay. However, these results could be due to feelings of affiliation with the agent as opposed to the apologies offered by the agent. The ability to project emotional expressiveness gives the impression of a more trustworthy, charismatic and credible individual. This is aided by others paying more attention to behavioural cues in order to obtain feedback on the progress of the situation (Lisetti and Nasoz, 2002). This property could be exploited in trying to design virtual agents in an e-commerce setting where trust and credibility play an important role.

A distinct characteristic in both studies, reported by Meehan (2001) and Prendinger et al. (2005), was the use of the individual's physiological responses in detecting stress. Although the analytical method used to derive results were basic through the admission of the authors, the usage of objective means of stress measurement is interesting. However, this suggests that even minimally responsive agents can have a significant impact on participants' physiological responses. However, many physiological measures may need to be used in tandem to understand the analysis of data gathered from participants exposed to affective virtual humans.

2.6 Quantifying success in virtual environments: Presence

The success of a VE is often measured in terms of the extent to which sensory data projected within a VE replaces the sensory data from the physical world (Sanchez-Vives and Slater, 2005). This success is conventionally quantified by rating the individual's sense of presence during the experience. Presence is partly to do with the VR technology and partly to do with the individual's state of mind (Schroeder, 2002). Presence, within the context of VEs, has been defined as the sense of '*being there*' in the place portrayed. Lombard and Ditton (1997) defined presence as the '*illusion of nonmediation*' and conceptualised presence under various viewpoints. Two of these concepts are of particular relevance to this research: 'social actor within the medium' and 'medium as a social actor'. The first concept of presence is to do with the extent to which individuals respond to behavioural and social cues portrayed by agents while the second concept is to do with the extent to which individuals treat an agent as a social being (Lombard and Ditton, 1997).

Numerous other definitions and conceptualisations of presence, copresence, social presence, spatial presence and temporal presence exist and can be found in (Zhao, 2003). Of all these concepts, copresence is the most relevant. Copresence generally refers to the individual's sense of being in the company of others. In this thesis, the more recent *operational* definition of presence and copresence as presented by (Sanchez-Vives and Slater, 2005) is used. According to this definition, presence is taken as the extent to which participants act and respond to virtual sense data as if it were real, where 'response' is considered at many different levels ranging from physiological through to cognitive. Similarly, copresence is taken as the extent to which participants act and respond to the agents as if they were real. In keeping with this view, if the VE and it's inhabitants (virtual humans) are successful in portraying a convincing scenario, it is expected that a participant's response to the virtually generated information is close to what could be expected under similar circumstances in the real world.

Traditionally, a participant's presence response is measured using various questionnaires including the Slater-Usoh-Steed (SUS) questionnaire (Slater et al., 1998; Slater and Steed, 2000), the Witmer and Singer (1998) presence questionnaire (PQ) and the ITC[7]-Sense of Presence Inventory (ITC-SOPI) (Lessiter et al., 2001). There has also been a lot of debate amongst researchers on deciding which questionnaire best measured presence in VEs (Usoh et al., 2000; Slater, 1999). Each questionnaire focuses on distinct areas of the experience which might contribute to presence in VEs. For instance, the SUS questionnaire focuses on the participant's sense of being in the place depicted by the VE, the participant's subjective feeling of the extent to which the VE becomes the reality, and the extent to which the experience is remembered as a place visited instead of images seen. These categories are similar to what have been found in factor-analytic studies e.g. (Lessiter et al., 2001; Schubert et al., 2001).

Recently, this method of collecting presence responses has been criticised due to its dependencies on the participant's accurate post-hoc recall, processing and rationalisation of their experience in the VE and their interpretation of the word 'presence' (Slater, 2004). Freeman et al. (1999) demonstrated that simple rating scales can be affected by prior experiences and called for the development of more structured questionnaires which can be used as reliable measures of reported presence in conjunction with other objective measures. In addition, Slater (2004) warns of the pitfalls in analysing collected data and drawing conclusions based on statistically significant results with no logical meaning.

Many objective means of measuring participant responses during the virtual experience have been suggested. These include a count of the number of *transitions* a participant experiences between the VE and the physical world termed breaks-in-presence (BIPs). Generally a higher sense of reported presence should tally with a lower number of reported BIPs (Brogni et al., 2003; Slater and Steed, 2000; Slater et al., 2003, 2006). Other measures of presence include psychophysiological responses (Meehan et al., 2002; Slater et al., 2006) and behavioural measures such as postural shifts or observations of social and physical behaviour (Bailenson et al., 2003; Freeman et al., 2000; Held and Durlach, 1992). In this

[7]ITC: Independent Television Commission

thesis, the premise is that a higher reported presence response will be associated with agents possessing expressive behaviours. Participant responses to the virtual humans were collected using a mixture of objective and subjective measures (Chapter 3).

2.7 Chapter Summary

The goal of this thesis is to model a set of affective behaviours (kinesics) in virtual humans in order to invoke realistic participant responses. This chapter has reviewed the theoretical basis of emotions and the complexities involved in creating affective virtual humans. Section 2.1 discussed the potential benefits to building expression into virtual humans while the importance of nonverbal behaviours in affective communication was reviewed in Section 2.2. The remaining sections dealt with existing computational models used in building expressive virtual humans and participant responses to affective expressions.

Research into building expressive virtual humans can be split into two themes: *generation* (Section 2.3) and the *representation* of emotional states (Section 2.4). A number of models for generating a virtual human's internal emotional state at run-time have been proposed. These models are especially useful in applications that require long-termed participant-virtual human interactions such as virtual therapy. However, as mentioned already, these emotion *generation* models can not be utilised effectively without models for *representing* the generated emotional states through appropriate behavioural cues. Unfortunately the later of the two research themes have been biased towards studying facial expression (Section 2.4.1). In particular, the role of kinesics (posture and body movement) in the communication of emotional states remains relatively unexplored (de Gelder, 2006).

Kinesics play an important role in the communication of affect (Sections 2.4.3.1 and 2.4.3.2). It has been argued that postural attributes and body movement may play a significant role in portraying an emotional state in conditions where the facial expression of the virtual human is obscured or when the virtual human is at a distance. Furthermore, although the face is the most closely observed and most accurately interpreted modality of expression, congruent postural cues and body movement may significantly improve nonverbal communication especially in full-body virtual humans. A lack of consistency in the nonverbal behaviours portrayed by a virtual human can lead to adverse participant responses and also to inaccurate interpretation of the virtual human's underlying emotional state (Dittmann et al., 1965; de Gelder, 2006; Montepare et al., 1999; Planalp et al., 1996). Yet, despite the extensive research conducted into expressive animation techniques, there are no parametric models of affective kinesic behaviour. Due to the degrees of freedom afforded by the human body, most models of affective body animation have included collecting motion captured data specific to a certain scenario and manipulating these to produce virtual scenarios. Unfortunately data-driven models are often cumbersome and inflexible. On the other hand, there has been little research into defining specific kinesic attributes that can be associated with a specific emotional state. This presents the main challenge in building a parametric model of affective kinesic behaviour.

Investigation into the behavioural cues used to express internal states can be used to build effective and economical parametric models as exemplified by the extensive research into the role of gaze in face-to-face communication (Section 2.4.2). Many gaze behavioural models were discussed in Section

2.4.2, two of which (Garau et al. (2001) and Lee et al. (2002)) were used to build a parametric model of gaze behaviour in a preliminary experiment (Section 4.1 of Chapter 4). The resultant gaze model was used to investigate the importance of modelling even minimal behavioural cues in full-body virtual humans. The same method is then applied to building parametric models of affective kinesic behaviours. Regrettably, most existing descriptions of affective displays of the body are confusing and not suitable for computation, however, Coulson (2004) provided a set of affective postures which were used as a starting point to investigate the role of postures in the communication of affect in full-body virtual humans (Chapter 5). A parametric model of kinesic behaviour for two emotional states (Anger and Sadness) was then developed for exploring body movement in Chapter 6.

Individuals in a private and highly emotional dyad arrange their postures in a vis-à-vis arrangement (Section 2.4.3.3). This was incorporated into the design of virtual scenarios for the evaluative experiments reported in this thesis. The main problem faced while evaluating affective virtual humans, objectively, is that there is no benchmark to compare the participant's response to the virtual human with their response to a real person in the same context. Therefore accurately interpreting the objective responses presents challenges. Subjective presence responses are collected using the SUS, ITC-SOPI and a questionnaire designed by Slater based on the operational definition of presence/copresence. In addition, objective and behavioural measures of participant responses were collected for the experiments designed to ascertain the role of kinesics (Chapters 5 and 6) in the portrayal of affect in virtual humans. The information reviewed in Section 2.5 was then used to interpret the results of the experiments.

Parts of this chapter are available in a workshop paper dealing with the issue of maintaining consistency between visual appearance and behavioural fidelity (Vinayagamoorthy et al., 2005) and a state of the art research paper dealing with the complexities of building plausible expression into virtual humans (Vinayagamoorthy et al., 2006b). The next chapter reviews the methods and materials used in the experiments on the nonverbal behaviour of virtual humans and participant responses. The experiments are described in more detail in Chapters 4, 5 and 6.

Chapter 3

Methods and Materials

This chapter discusses the methods, materials and techniques used in the experiments discussed in Chapters 4, 5 and 6. Each experiment is slightly different and these differences are elaborated in the corresponding chapters.

A pure-experimental approach combined with primarily quantitative data acquisition techniques was adopted to investigate the research questions in this thesis due to its potential power to discern cause-effect relationships with more confidence. In addition, some qualitative methods were utilised with the view to explaining participant responses to the virtual humans in the experiment. Generally, the experimental approach has been used hand in hand with quantitative methods of acquiring and analysing data especially within the realms of virtual reality. However, there are potential advantages to using the qualitative analysis of post-experimental interview data with the view to uncover new avenues for research (Garau, 2003). Furthermore, using a combination of methods in collecting participant responses has been successful in many other fields of research (Tashakkori and Teddlie, 1998; Creswell, 2003). Combining a controlled laboratory-based experimental approach with various data-collection methods provided a number of advantages. These include reinforcement of any results with the addition of a reason to any quantitative results, the possibility of revealing a reason for contradictory results, the potential discovery of new avenues for research and validating the aforementioned advantages of using a combined research methodology with respect to the study of virtual humans.

Section 3.1 reviews a brief critique of some of the different methods used in designing experiments and respective data acquisition techniques. Section 3.2 focuses on the design of the experiments reported in this thesis. Section 3.3 deals with the methods used to collect data and participant responses while Section 3.4 focuses on the methods used to extract results from the data. Finally, a summary of the chapter is presented in Section 3.5.

3.1 Research choices: qualitative and quantitative approaches

The longstanding debate on the merits of the use of the qualitative versus quantitative approaches is often referred to as the "qualitative-quantitative" debate (Glassner and Moreno, 1989; Creswell, 2003). In general, qualitative research focuses on generating rich and detailed data that can contribute to an in-depth understanding of the context with respect to classification and involves asking definitive questions such as 'what?' and 'why?' about the collected data. On the other hand, quantitative research generates population-based and generalisable data based on a pre-defined hypothesis. This is well suited to establishing cause-and-effect relationships to answer enumerative type questions of 'how many?' and 'what is the strength of association?'. Most surveys, case and field studies, follow the qualitative approach while laboratory-based controlled experiments ascribe to the quantitative methods. Qualitative methods are often thought of as lacking in reliability since the interpretation of results often depends on the analyst while quantitative methods are criticised because of their need for a controlled setting and lack of ecological validity (Glassner and Moreno, 1989; Creswell, 2003). The philosophical standoff between subscribers to either paradigm is generally irreconcilable due to the vast differences between them (Table 3.1). As Glassner and Moreno (1989) put it, *"the quantitative scientist reduces words while the qualitative scientist enlarges them"*.

	Quantitative (*top-down*)	**Qualitative (*bottom-up*)**	**Mixed**
Method	Objective	Subjective	Both
Focus	Narrow	Wide	Multi
Reasoning	Deductive (logical)	Inductive	Both
Setting	Controlled (experimental/closed)	Flexible (natural/open)	Controlled
Hypotheses	Testing (Predicted)	Generation (Interpretation)	Testing
Sampling	Statistical (predetermined)	Purposive (evolving)	Statistical
Measurement	Logging instruments	Researcher as instrument	Both
Data Coding	Numerical (statistical)	Words/Categories/Themes	Both
Data Analysis	Statistical inference	Categorising/Comparing	Both
Final Report	Statistical	Narrative	Pragmatic

Table 3.1: Comparisons between the qualitative, quantitative and the mixed method used in this thesis

Despite these differences a pragmatic solution is often chosen, whereby the researcher weighs the strengths and weaknesses of both paradigms and chooses a mixture of methods in accordance to the design of the particular study: a *mixed* method approach. Tashakkori and Teddlie (1998) presented Greene et al.'s (1989) five purposes for adapting the mixed methods approach; triangulation (seeks convergence of results), complementarity (examine overlapping and different facets), initiation (discover paradoxes, contradictions, fresh perspectives), sequential development (results from one method shapes the other) and expansion (adds richer details to the results). A mixed method approach employs strategies of inquiry that involve the collection of data in both numerical and textual forms so that the final database consists of both qualitative and quantitative information which are analysed together to form a coherent interpretation of results.

The strategy adopted in this thesis is the *concurrent triangulation* method which is characterised by an *initial* phase involving the collection of both quantitative and qualitative data concurrently. This is followed by an *interpretation* phase. Ideally both types of data are analysed at the same time but in this thesis, priority was given to the quantitative data due to the added advantage offered in terms of falsifiability. Results were integrated during the interpretation phase in an attempt to uncover any convergence of both sets of results. This strategy was chosen since it could potentially lead to well-validated and substantiated findings (Creswell, 2003). In addition, the data was collected in a shorter time span in comparison to sequential mixed methods. There were limitations in following the concurrent path in that it required careful consideration while studying and comparing the same phenomenon in two separate forms especially in the case of discrepancies. This method also called for a set of research questions that optimally fitted both methods. The following section outlines the experimental design, apparatus, procedures and data collection techniques used in the experiment while Section 3.4 covers methods used to analyse the data.

3.2 The Experiments

Although the experimental approach has been chosen for this thesis, the hypothesis were generated based on theories that were a result of observational studies in the fields of social psychology and behavioural sciences. The main experiments described in Chapters 5 and 6 were designed based on experience and results from the preliminary experiments reported in Chapter 4. In the preliminary experiments, designs were drawn with the view to explore the importance of behaviours in full-body virtual humans in IVEs. The main experiments were then designed to investigate the role of specific behavioural cues (***kinesics***) in the portrayal of two specific emotional states (***Anger*** and *Sadness*).

3.2.1 Designing for the research question: Variables

In designing experiments it is important to address the research question at hand by defining appropriate *dependent*, *independent*, and *explanatory*.

Independent variables are meant to cause and influence differences, if any, in the outcome of dependent variables. The independent variables in this thesis were defined based on aspects of the virtual human's behavioural capability. The independent variables were manipulated to create different conditions within the experiments while dependent variables were recorded using a combination of subjective questionnaires, post-experiential interviews, behaviour observations and psychophysiological data. The values of the dependent variables were then analysed in conjunction with explanatory variables to extract effects due to independent variables.

The explanatory variables in an experiment were the variables considered in the statistical analysis in order to eliminate their influence on the dependent variables (participant responses). This ensured that the true influence of the independent variable on the dependent variable became apparent. For instance, Freeman et al. (2003) reported that participant prone to high levels of anxiety in the physical world were susceptible to paranoia towards agents. In this thesis, the explanatory variables were recorded in the form of a demographic pre-experiential questionnaire data, the participant's traits in accordance to various standard psychological questionnaires and an objective baseline measure of the participant's physiological state. The physiological baseline was used to root out variation in the data which could be due to individual differences in the participant's physiology.

3.2.2 The design

Creswell (2003) categorised experimental designs into four main groups; *pre-experimental* (single group), *quasi-experimental* (participants are not randomly assigned), *true* experiment (randomly assigned participants) and *single-subject* (observing a single or small number of participants over time).

All the experiments reported in Chapters 4 and 5 are of a between-group true experimental design. In each experiment, a factorial design with two independent variables was employed. Each *factor* in the experiment equated to an independent variable being tested and could be manipulated to create four different *levels* or *conditions (or cells)* in the experiment. The data collected in these conditions were analysed to determine the independent and interactive influence of the factors on the dependent variable

		Factor 1	
		Level 1	*Level 2*
Factor 2	*Level 1*	Condition 1	Condition 2
	Level 2	Condition 3	Condition 4

Table 3.2: The 2x2 factorial design

The main advantage of this experimental design is that it permits the study of any significant interaction effects between the independent variables. The between-group design also ensured each participant was kept completely unaware of the other conditions in the experiment hence the participant was unable to form expectations about the virtual human in a particular condition and bias responses which is often a problematic issue in a within-group experiment. The between-group design favours experiments which try to uncover the relative importance of the independent variables being measured. For instance, the first preliminary experiment was designed to explore two aspects of virtual human fidelity: visual appearance versus behavioural fidelity (gaze). However, the lack of comparisons between conditions in the between-group design can also be viewed as a disadvantage since participants in factorial experiments can not be asked to draw on experience to form evaluative judgements of the condition. This is especially troublesome in VR-based experiments as participants are unlikely to have experienced IVEs prior to the experiment. Another disadvantage in a between-group experiment is the increased number of participants needed in comparison to a within-group design.

Factor 1
Condition 1
Condition 2
Condition 3

Table 3.3: The within-group design

The final experiment, reported in Chapter 6, was based on a within-group design with three levels. Each participant was asked to experience all three levels of the experiments. This gives the researcher an opportunity to investigate the differences between participant responses to all the levels. The advantage of this method is that a small number of participants yield a rich amount of data. In the final experiment, where a deeper understanding of participant responses was needed, this design was more conducive to revealing a more coherent story. The main disadvantage of the within-group design was that participants developed expectations of the other levels in the experiment once they have been exposed to the first. This effect was minimised by randomising the order in which the conditions were presented to the participant. This ensured that results due to the order of presentation was uncovered during the analysis.

3.2.3 Apparatus

This section gives an overview of all the systems and devices used in the experiments presented in this thesis. Slight variations of the apparatus were made in order to accommodate the design of each experiment. These variations are discussed in the chapters corresponding to the experiments.

The apparatus described in Sections 3.2.3.1 and 3.2.3.2 are virtual reality (VR) systems that allowed the participants to move while remaining fully immersed in the virtual environment displayed and view it in the correct perspective. The first is a VR system, the Trimension ReaCTor, in which high resolution images are presented to the participant on three surrounding walls and a floor, with continuous reference to their viewpoint updated via a head-tracking device. The second, a Head Mounted Display (HMD), is a VR device that participants wear on the head to have visual information directly displayed in front of their eyes. In both systems, images are presented to participants with a separate left and right image to their left and right eye respectively. This gives the illusion of a perceived 3D space within and beyond the actual space. The advantages of using a fully *immersive* VR system includes a wide field of view, full immersion of the participant's body within the virtual space and full-life scale visualisations of virtual models including humans.

The Trimension ReaCTor was used in all the experiments. In the first preliminary experiment (Chapter 4), the HMD was used, in addition to the Trimension ReaCTor. The ProComp+ device discussed in Section 3.2.3.3 was used to record participants' physiological responses to events and stimuli in the IVEs.

3.2.3.1 The Trimension ReaCTor

The Trimension ReaCTor, a CAVE^{TM}-like system (Cruz-Neira et al., 1993). It consisted of three 3m x 2.2m walls and a 3m x 3m floor as depicted in Figure 3.1. The front, left and right walls were made of acrylic screens on which images were back-projected while corresponding images were projected on the white painted wooden floor from above. The system was powered by a Silicon Graphics Onyx2 with 8 300MHz R12000 MIPS processors, 8GB RAM and 4 Infinite Reality2 graphics pipes. The participants wore lightweight CrystalEyes stereo glasses, which were tracked by an Intersense IS-900 system with 6 degree of freedom (DOF) inertial motion tracking system accurate to within 2mm with an end-to-end latency of 50ms. The inertial components were used to track the position and orientation of the participant while ultrasonic sensors were used to correct drifts in the inertial sensors. The participants used a tracked wand with five buttons and a mini-joystick to navigate the VE. This wand was used in the first preliminary experiment discussed in Section 4.1. It was replaced with a similar but *wireless* wand in all the other experiments. Ordinarily the five buttons could be assigned various functionalities that could be used to interact with the VE such as grasping and rotating objects. The five buttons were

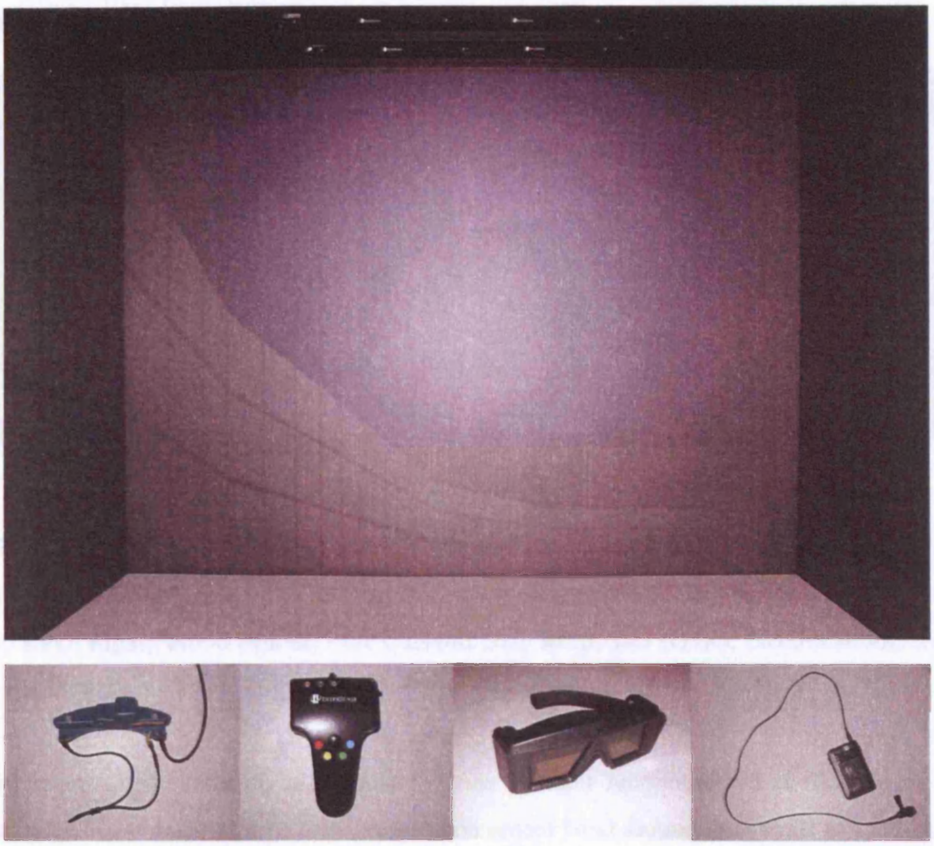


Figure 3.1: The Trimension ReaCTor: a CAVE^{TM}-like system used; Bottom (Left to right) - tethered head tracker, tracked wireless wand, stereo glasses and wireless microphone.

disabled in all the experiments reported within this thesis by making all virtual objects non-graspable. The mini-joystick was used for navigating the VE, with the pointing direction determining the direction of movement on the horizontal plane. The speed was controllable to a small extent by the amount of force exerted on the joystick. The participant could also move naturally within the physical constraints of the three walls and floor.

3.2.3.2 The Head Mounted Display (HMD)

The Head Mounted Display (HMD) used was a Virtual Research V8 with a true VGA resolution of 640x480x3 colour elements for each eye. The scenarios were implemented on a Silicon Graphics Onyx with twin 196 MHz R10000, Infinite Reality Graphics and 192M main memory. The tracking system had two Polhemus Fastraks, one for the HMD and another for a 5 button 3D mouse (wand). Similar to the ReaCTor, the only control on the wand enabled was the joystick used to navigate the VE. The V8 had a field of view of 60°diagonal at 100% overlap. The disadvantage of using a HMD in comparison to the Trimension ReaCTor was the increased amount of hardware on the participant. Similar to the ReaCTor, participants could move within the VE using either the wand or physical movements. However, physically moving, while wearing the HMD, was more difficult since participants could not see the physical world at the same time.

3.2.3.3 The Physiological monitoring device: ProComp+

(a) (b)

Figure 3.2: a) A participant kitted out with the physiological sensors in the ReaCTor; b) The Thought Technologies ProComp+ device, (inset) an illustration of the optical fibre cable used to connect the device to a PC; Right (Top to bottom) - the Galvanic Skin Responses (GSR), Electrocardiogram (ECG) and the respiration sensors.

Participants were fitted with an 8 channel device Thought Technologies Ltd. ProComp+ Electrocardiogram (ECG) sensors on their torso, respiration sensor band around their chest and Galvanic Skin Response (GSR) sensors on their non-dominant hand. The ProComp+ Infinity device records the ECG at 256 Hz and GSR and respiration at 32 Hz (Figure 3.2(b)). The device is connected to a PC through a optical fibre cable linked to a connector which fits into the serial port of a PC. It is linked to the virtual reality system using a Virtual Reality Peripheral Network (VRPN) driver, that allows access through the network to the data streaming, coming both from the physiological sensors and the tracking sensors. Figure 3.2(a) shows a participant wearing the tracking and physiological sensors while viewing the training room on the right hand side wall of the ReaCTor. Figure D.1, in Appendix D, depicts a typical setup of the ReaCTor and the physiological device in usage for the experiment discussed in Chapter 5.

3.2.4 Software

The main software used to implement the experiments presented in this thesis was Distributed Interactive Virtual Environment (DIVE) (SICS, 2006). This software was augmented with two main plugins, Virtual Reality Peripheral Network (VRPN) and Platform Independent Architecture for Virtual Character and Avatars (PIAVCA), in all the experiments other than the first preliminary experiment (Section 4.1 of Chapter 4).

3.2.4.1 Distributed Interactive Virtual Environment

The virtual humans and environments were implemented on a derivative of Distributed Interactive Virtual Environment (DIVE) 3.3x that was extended to support spatially immersive systems (Frécon et al., 2001; Steed et al., 2001). DIVE is an internet-based multi-user VR system in which participants can navigate in a shared 3D space and interact with each other. Every DIVE application at a networked site keeps its own copy of the world database. When the state of an object changes in one DIVE application, the change is distributed to all connected applications. This property is used to enable participants to collaborate within a shared VE. Each object within a DIVE world can be associated with a number of properties which can be used to change the state of the object. For instance objects can easily be set to an invisible status, made non-graspable and set to detect collision with the avatar of a participant.

Participants at each site in a DIVE world, are represented by an avatar (DIVE actor). Mortensen et al. (2002) successfully used DIVE to represent participants on separate continents within a shared VE over Internet2. DIVE reads the participant's input devices (head trackers and wands) and maps the physical actions taken by the participant to logical actions in the DIVE system. This includes navigation in 3D space, clicking on objects, grabbing objects and so on. DIVE also supports the import and export of virtual humans, objects and environments in VRML and several other 3D file formats. In addition, DIVE is made modular through C/C++ plugins, which act as a dynamically loadable library of symbols and functions. The benefits of this facility include simpler reconfiguration of the software and the integration of new technologies. This facility was used extensively in the creation of animations for all the virtual humans used in the experiment.

A large part of the functionality of DIVE are also accessible using it's Tcl interface. The DIVE/TCL interface accepts normal Tcl commands and defines a set of DIVE interface functions. The dynamic behaviour of virtual objects, such as automatic opening of virtual doors, were described by interpretative Tcl scripts.

3.2.4.2 Virtual Reality Peripheral Network

Virtual-Reality Peripheral Network (VRPN) is a public-domain, open-source software which provides a device-independent and network-transparent interface to VR peripherals. It is a set of classes within a library and a set of servers designed to implement an interface between VR application programs and a set of physical devices (tracker, wand, physiology recording devices etc.) (Taylor II et al., 2001). VRPN allows for simultaneous connections between devices on different machines and allows for synchronised logging of time-stamped data. The VRPN software used to record physiological data had two simultaneously running programs: a client and a server. The server communicated with the ProComp+ device and time-stamped the incoming data. It was also used to record DIVE events which were sent from the virtual environment to annotate participant-triggered events such as opening of doors and interactions with the agents. The server forwarded the data to the client, which recorded and displayed a graph of the data at run-time.

3.2.4.3 Platform Independent Architecture for Virtual Character and Avatars

The virtual humans used in the second preliminary experiment (Section 4.2) and the two main experiments (Chapters 5 and 6), were animated using a version of Gillies et al.'s (2005) Platform Independent Architecture for Virtual Character and Avatars (PIAVCA).

PIAVCA is a virtual human animation library that is designed to be independent of any underlying graphics engine, and has been implemented as a plugin to DIVE (Section 3.2.4.1). There are four key features to PIAVCA which made it especially suited to the experiments. Firstly PIAVCA abstracts away from the details of any underlying graphics engine so the experiment scenarios were built independent of any underlying platform. Each component in PIAVCA is divided into two components; a front end independent of the underlying graphics engine and a back end which contains all methods specific to the platform (in this case DIVE). There are two main components to PIAVCA: *avatars* and *motions*. PIAVCA avatars (including agents) can be animated by directly moving the body parts of the agent, or through a motion loaded into the PIAVCA avatar.

Secondly PIAVCA features a unified abstraction for different styles of animation including key-frame animation and procedural animation. PIAVCA also provides a variety of motion filters and combiners which were used to transform combine a basic set of motions to produce a much larger set of animations. This allowed for the creation of a wide range of animations and behaviours at run-time using a relatively limited set of atomic motions. Motion filters and combiners allow the experimenter to manipulate motions in a number of ways. Motion files could be played on a repeated loop, with a slight delay, blended with another motion in a variety of ways or a combination of all of the above. Motion files could also be played in a manner that queued one motion after the other so that the last frame of the first animation blended smoothly into the first frame of the next animation. For instance, the animations which allowed the agents to turn towards participants in the experiments discussed in Chapters 5 and 6 were created at run-time by blending a base head posture with two 90 degree head animations. The resultant animations were then placed on a queue to create smooth head turning animations. In addition to playing motion files, PIAVCA supports the addition of background animations to a virtual human. This is akin to perlin noise (Perlin and Goldberg, 1996), however, the background animation in PIAVCA could be a simple breathing or blinking animation as well. This functionality was used to make the agents appear less robotic when no animation files were being played.

Thirdly, PIAVCA supports facial mesh animations. All motion combining and filtering functions described for the skeletal/postural animation of the virtual humans were made available to facial mesh animations as well. The facial expressions are created in 3D Studio Max (Autodesk, 2006) and the mesh deformations of the agent's faces were directly imported into PIAVCA to produce appropriate facial cues. Finally PIAVCA supports the use of Biovision motion (BVH[1]) file format and allowed for the creation of extensions to it's core functionality. BVH is a file format that can be used to represent human motion in a form that is independent of the virtual human. This enabled a single piece of motion to be applied to multiple virtual humans. All postures and animation files used in the experiments were created in 3D Studio Max (Autodesk, 2006) and converted to the BVH format using Nugraf (Okino, 2006). Like DIVE, PIAVCA also allows for the creation of new modules in C/C++ as an extension to the core PIAVCA application. Extension made to PIAVCA are discussed in the chapters corresponding to the appropriate experiment.

3.2.5 Ethical considerations

All experiments reported in this thesis were pre-approved by University College London's Committee on the Ethics of Non-NHS Human Research and required registration with the Data Protection Officer. The applications were made at least two months before the study commenced.

[1] BVH stands for Biovision hierarchical data

The main elements that were required for good ethical conduct included:

- Giving the participant sufficient information about the experiment (and potential risks) so that each participant was able to take an informed decision about their participation. The participant is also informed of their right to ask questions.
- Each participant was required to sign a consent form confirming their approval in taking part in the experiment and releasing collected data for use in the study.
- All participants were informed of their right to withdraw from the experiment at any time during the experiment without the need to give a reason.
- Anonymity is an important part of the ethical guidelines and every precaution was taken in order to protect data that could be traced back to any particular participant. For instance, each participant was give a numerical ID which differentiated their data set from others and all findings are reported in aggregate.
- All personal data collected about the participant including the consent forms were kept in the strictest confidence.
- Participants were not deceived about the procedures they had to undertake in the experiment, however, some experiments such as the ones reported in this thesis require that the participant was steered away from the actual aim of the experiment. However, this false pretence was explained to the participant in the debrief sessions at the end.
- All participants involved in experiments reported in this thesis were paid compensation at the rate of £5.00 per hour. They were also given the experimenter's contact information.

3.2.6 Piloting

A sample of participants were invited to take part in early versions of each experiment. Informal observations and detailed feedback from these participants helped to improve some of the questionnaires, the task, and the design of the experiment and the virtual environment. The pilots also acted as rehearsals to familiarise the experimenters and/or confederates on the procedures. This ensured the smooth running of the experiments and to some extent helped identify more confounding variables. None of the data collected in the pilots was used in the final analysis.

3.2.7 Procedure

With the exception of the experiment reported in Chapter 6, all experiments were conducted with the help of colleagues. The reason for this was two-fold. Firstly the procedures were too complicated and required two experimenters to enable the smooth running of the experiments. Secondly the experiments were designed around having two experimenters. For instance, it allowed the experimenters to give vital information and task-related instructions twice to each participant. The roles played by the experimenters were also randomised with a view to minimise possible effects of confounding variables. Even though different experimenters were involved, the procedures were standardised across all the experiments through running sufficient pilots. In the case of the main experiment run by only one experimenter, the procedures were slightly different and less complex (Chapter 6).

The participants were recruited through poster, email and/or online campaigns through the university campus. They were emailed information detailing some of the procedures involved and the apparatus used in the experiment. Preliminary consent was obtained through email to ensure that a number of criteria were fulfilled. For instance individuals who had a history of epilepsy were screened since there was a chance that the VR apparatus could trigger an epileptic episode. In the main experiments (Chapters 5 and 6), female participants were excluded. Once the potential participant gave partial consent to taking part in the study, a time slot was negotiated with them and the preset identity code was assigned to that participant. Directions to the laboratory were sent via email and repeated reminders were sent as the day of the experiments approached in order to ensure the participant either showed up or had ample opportunity to withdraw.

On the day of the experiment, the experimenters always checked the apparatus used to ensure it was in working order. Once the participant arrived, the experimenters welcomed the participant in an anteroom and give an information sheet explaining the procedure, the apparatus and possible risks involved and the task (Appendices E.2 and G.5). The participant was encouraged to ask questions. The participant was informed that they could withdraw from the experiment at any time without obligations, that all data collected would be kept confidential and any findings reported would be in aggregate. They were then asked to sign a consent form asking for a number of things including their consent to take part in the experiment in addition to being audio taped, video taped and physiologically monitored (Appendices E.3 and G.6). The participant was then allowed to fill in some pre-experimental questionnaires on their own. The questionnaires were designed to collect demographic and personal data as explanatory variables. These questionnaires varied with the experiment and are discussed in the appropriate chapters.

The participant was then invited into the appropriate VR system used in the experiment. Following this, the experimenters fitted the participant with physiological monitoring devices and tracking sensors. In the past, participants have been left to attach the sensors by themselves in the interest of minimising undue stress, however, this has lead to the collection of unusable data due to the participant's unfamiliarity with the apparatus. No physiological data was collected in the first preliminary experiment (Section 4.1 of Chapter 4). The participant was then asked to remain in a relaxed state (still and quiet) for a few minutes in order to record baselines for the physiological measurements. Again the baseline step was not carried out for the first preliminary experiment. In the final experiment, the baseline step was carried out four times (Chapter 6). Afterwards the participant was given adequate training on navigating VEs and reminded of the task. The virtual training environment was separated from the actual virtual environment used in the experiment and less visually stimulating. In addition to providing an introduction to the virtual space, the training process acted as a screening process as well. Participants prone to simulator sickness would normally be weeded out during the training session especially since the training routine generally required more focus and movement than the actual experiment. During the training session, the participant was introduced to the various aspects of an IVE. For instance, a large number of participants were unaware of the ability to use natural movements (such as squatting, looking around objects, or turning the head) to explore the VE.

Most experiments had a time limit within which the participant was asked to complete the task in order to standardise the experience of all participants across the experiment. Another reason for setting this time limit was to minimise the possibility of getting simulator sickness during the experiment. The experimenters always reminded the participant of their right to withdraw from the experiment at any time. While the participant carried out the task in the VE, the experimenters quietly recorded their observations of the participant. On completing the task (or reaching the end of their time limit), the participant was administered post-experimental questionnaires after which the experimenters conducted a semi-structured interview. Interviews were only conducted in the second preliminary experiment (Section 4.2 of Chapter 4) and first main experiment (Chapter 5). Finally the participant was debriefed about the true goals of the research, paid and asked not to discuss the experiment with others who might be prospective participants.

3.3 Data Acquisition Techniques

Deciding on the appropriate data collection technique depends on the purpose of the experiment (Tashakkori and Teddlie, 1998). In the observational approach, data is collected from specific instances with the view to build an abstraction or theory therefore traditionally researchers focus on the design of qualitative data collection strategies. On the other hand, in an experimental approach, cause-and-effect relationships are predicted beforehand and data is collected with the view to either verify the hypothesis or falsify it. Within this approach, data collection calls for relatively focused, detailed and structured techniques and/or instruments. In essence, both the quantitative and qualitative traditions identify data acquisition techniques that may be arranged on a continuum: unplanned to highly planned (Tashakkori and Teddlie, 1998). Most of the data collection techniques used in the experiments contributing to this thesis were on the higher end of this continuum.

3.3.1 Subjective responses

Subjective measures are one of the most traditional ways of obtaining data in both qualitative and quantitative research. The subjective data acquisition techniques used in the experiments within this thesis were mostly close-ended questionnaires, Likert-type response scales and semi-structured interviews.

3.3.1.1 Close-ended questionnaires and Semi-structured interviews

Close-ended questions were used in the experiments since they provide a uniformity of answers and are easier to analyse. Most of the questionnaires used in the experiments consisted of dichotomous or Likert-type close-ended questions. In the experiments, the experimenter left the participant to complete the questionnaires. This minimised biased responses due to participant reactivity. However, the experimenters did not leave room for the collection of extra data on "interesting" cases. Also they required experimenters to postulate all the possible types of dependent variables before the experiment. In order to counter-balance for this, two of the experiments reported in this thesis also involved a semi-structured interview with the participant. This provided an opportunity to clarify responses.

In this research, unstructured interviews were used informally during pilots in an attempt to get an insight into how to improve the final version of the experiment. During the experiments, audio-taped semi-structured interviews were conducted. This offered the advantage of allowing the experimenter to conduct informal, conversational-like in-depth interview with participants while adhering to some level of structure. Each interview in the experiment started with a relatively general but experiment-related descriptive question (ice-breakers) to help establish rapport with the participants. There was a pre-planned agenda but this only listed the general topics to be covered in the interview thereby making it slightly easier to code the data collected. A general framework of open-ended evaluative questions with room for follow-up questions was planned beforehand and followed through in no particular order during the interviews (Appendix E.10). Appropriate prompts and probing questions were used when needed to elicit more in-depth feedback. The path followed and exact phrases used emerged through the course of each interview. However, these interviews were time-consuming and were only formally conducted in the experiments which were designed as exploratory studies (Chapters 4 and 5).

3.3.2 Objective measures

The participants' post-hoc rationalisations is one of the main disadvantages of using subjective measures of participant responses. In addition, it has been argued that the use of questionnaires to measure concepts such as presence is problematic since they are sensitive to prior experiences (Freeman et al., 2000), may not be able to distinguish between reality and virtual reality (Usoh et al., 2000) and are subject to the participants interpretation of the concept (Slater, 2004). Objective data, on the other hand, are 'external to the mind', concern facts and the precise measurement of concepts. For instance, internal physiological data is generally reliable but it is harder to measure, code and interpret. Recently, Slater (2004) argued that a scientific basis for presence in VEs cannot be established on the basis of post-experiential questionnaires alone. Psychophysiological and other objective data can be used to rate the effectiveness of the VE in addition to subjective measures. Meehan et al. (2002) used physiological reactions to a virtual precipice as an objective measure of presence in VEs, however, it is not known if physiological responses to less startling social situations in virtual environments are as detectable.

This is not to state that objective data is better or more valuable than subjective data. Research using physiological and other behavioural measures as an objective measure of success in VEs is still in its infancy. It is difficult to make inferences regarding a participant's intentions, motivations and attributes based on observed behaviours alone (Tashakkori and Teddlie, 1998). In this thesis, mixed methods of data acquisition were used in tandem with the view to produce complimentary findings.

3.3.2.1 Psychophysiological measures

The nervous system is a highly integrated structure but Figure 3.3 shows an overview of the nervous system. Most of the psychophysiological responses of interest are controlled by the Autonomic Nervous System (ANS) which control the visceral structures (glands and organs) of the body. The ANS is the regulator and coordinator of important bodily activities including body temperature, blood pressure and many aspects of emotional behaviour. Its main function is to keep a constant internal body environment by regulating bodily functions that could change as a result of internal or external stimuli. In extreme

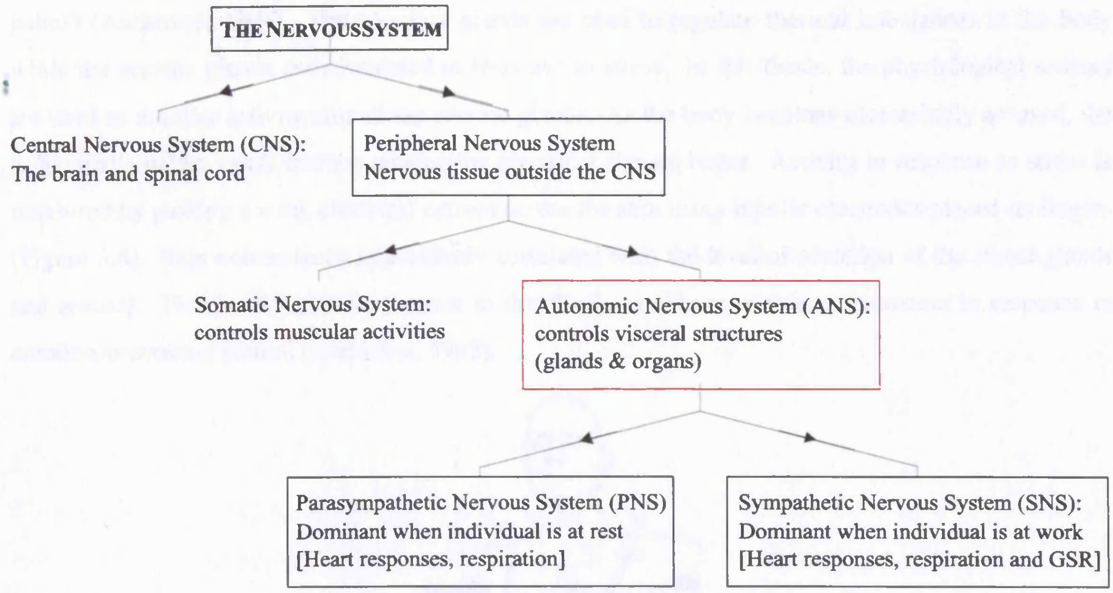


Figure 3.3: A brief structure of the nervous system as adapted from Andreassi (1995)

circumstances, the ANS mobilises the resources needed by the body to respond to situations of stress Andreassi (1995). For instance Meehan (2001) report that the heart rate can double in 3-5 seconds while sweating can begin within 2 seconds.

A wide range of physiological responses can be measured including galvanic skin responses (GSR), electrocardiogram (ECG), respiration, blood volume pulse (BVP), pupillometry (measure of changes in pupil size), electrooculography (EOG - measure of eye movement), electroencephalogram (EEG), electromyogram (EMG - a measure of muscle activity), salivation, skin temperature and so on (Andreassi, 1995). Three physiological responses were collected in the last three experiments reported in this thesis: GSR, ECG and respiration.

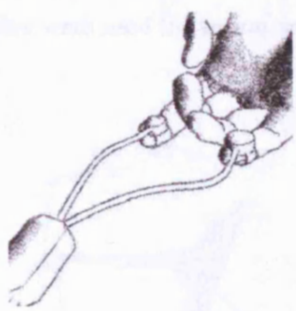


Figure 3.4: Placement of the GSR sensors

Galvanic Skin Response (GSR) or Electrodermal Activity (EDA) is measurable in terms of skin potential or skin conductance. GSR measures the phenomenon in human physiology during which the skin momentarily becomes a slightly better conductor of electricity due to an increase in sweat gland activity. This occurs when the individual is physiologically aroused (Scheirer et al., 2002). Sweat is produced by two types of glands in the human body: *apocrine* (e.g. the armpits) and *eccrine* (e.g. the

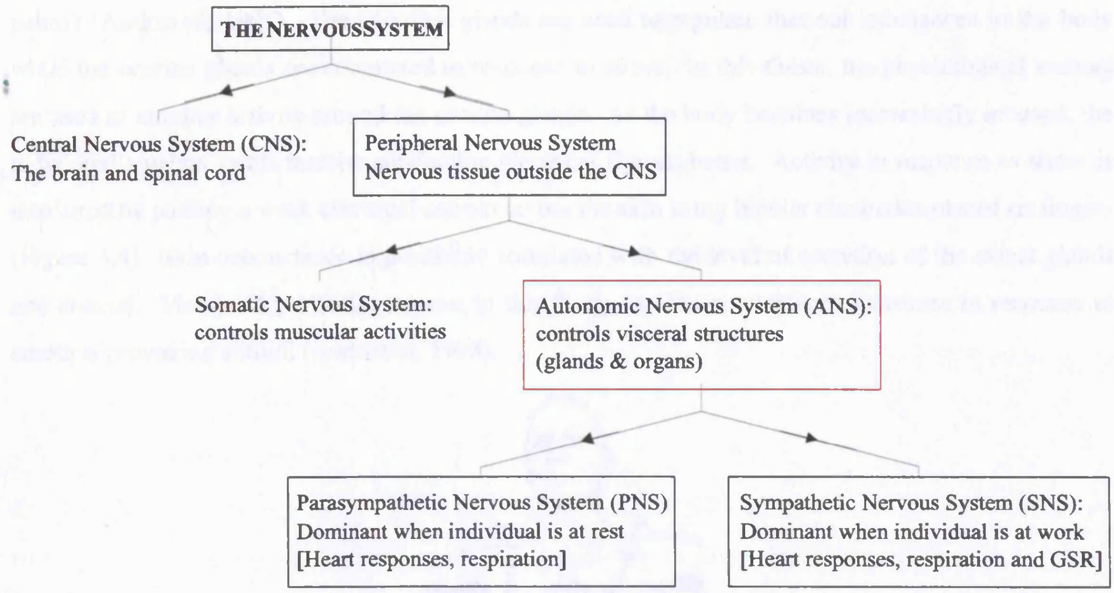


Figure 3.3: A brief structure of the nervous system as adapted from Andreassi (1995)

circumstances, the ANS mobilises the resources needed by the body to respond to situations of stress Andreassi (1995). For instance Meehan (2001) report that the heart rate can double in 3-5 seconds while sweating can begin within 2 seconds.

A wide range of physiological responses can be measured including galvanic skin responses (GSR), electrocardiogram (ECG), respiration, blood volume pulse (BVP), pupillometry (measure of changes in pupil size), electrooculography (EOG - measure of eye movement), electroencephalogram (EEG), electromyogram (EMG - a measure of muscle activity), salivation, skin temperature and so on (Andreassi, 1995). Three physiological responses were collected in the last three experiments reported in this thesis: GSR, ECG and respiration.

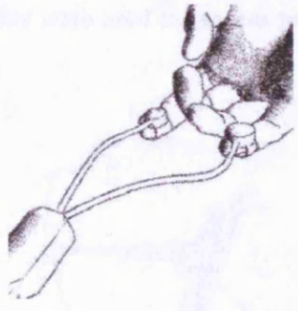


Figure 3.4: Placement of the GSR sensors

Galvanic Skin Response (GSR) or Electrodermal Activity (EDA) is measurable in terms of skin potential or skin conductance. GSR measures the phenomenon in human physiology during which the skin momentarily becomes a slightly better conductor of electricity due to an increase in sweat gland activity. This occurs when the individual is physiologically aroused (Scheirer et al., 2002). Sweat is produced by two types of glands in the human body: *apocrine* (e.g. the armpits) and *eccrine* (e.g. the palms) (Andreassi, 1995). The apocrine glands are used to regulate thermal imbalances in the body while the eccrine glands are stimulated in response to stress. In this thesis, the physiological sensors are used to monitor activity around the eccrine glands. As the body becomes increasingly aroused, the individual's palms sweat thereby conducting electrical signals better. Activity in response to stress is measured by passing a weak electrical current across the skin using bipolar electrodes placed on fingers (Figure 3.4). Skin conductance is positively correlated with the level of secretion of the sweat glands and arousal. This is of particular interest to this thesis as skin conductance increases in response to emotion-provoking stimuli (Andreassi, 1995).

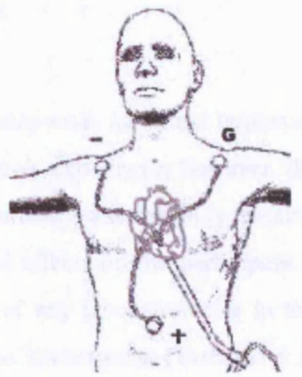


Figure 3.5: Placement of the ECG sensors

Heart responses to stimuli can be measured in many ways including an electrocardiogram (ECG/EKG) which is an electrical recording of the heartbeat. An enhanced heart rate reflects a reduced regulation of blood pressure resulting in enhanced blood pressure variability. Blood pressure is at its highest during a heartbeat. This is recorded as the *systolic pressure*. When the heart is between beats blood pressure falls. This is the *diastolic pressure*. An ECG measurement is recorded by monitoring the electrical changes that accompany every contraction of the heart. In the experiments reported in this thesis, three ECG measurement sensors were used in tandem to log the participant's heart rate in the experiments (Figure 3.5).

Figure 3.6: Placement of the respiration sensors

Respiration refers to the breathing rate of the participants. This was measured by wrapping an elastic band around the participant's chest (Figure 3.6). The sensors use the variations in the width of the rib cage to record the rate of respiration of participants.

These responses were normally recorded during different periods of the same experiment and then compared and contrasted to detect changes in the participant's physiology. For instance, in the second preliminary experiment presented in Chapter 4, these measures were collected during each study for three subsets of the duration of the experiment: the baseline period (1.5 minutes), the training period and the experiment (3 minutes each). Normal physiology levels often vary widely between participants so the analysis took into account each participant's baseline recording. This allowed for the interpretation of changes in the psychophysiology of participants. The physiological measurements were also recorded during the time used for training the participant on how to navigate in virtual environments, however, this data was not particularly analysed.

3.3.2.2 Observational records

Observational methods of gathering responses to virtual humans eliminate the need to depend on the participant for an accurate recall of their experience, however, there are other problems. A key issue in the case of observing or video recording participants is reactivity. An awareness of being observed might have other social-psychological effects on the participant. For instance, when an individual is being observed, they become aware of any inconsistencies in their behaviour which in turn leads to an altered behaviour in response to the environment (Tashakkori and Teddlie, 1998). Participants were observed by an experimenter who took notes on their behaviour in response to events in the VE. The participants were also recorded on video tape. These observational records were used as points of interest in interviews. For instance, in the second preliminary experiment (Section 4.2 of Chapter 4), participants were interviewed while they viewed the recorded video. In order to eliminate this dependency, record and replay softwares have been written to enable experimenters to experience the VE through the point of view of the participant's avatar, however, these types of software are still in the experimental stages of development (Steed et al., 2005).

3.3.2.3 Behavioural response: Participant-agents spacing

Bailenson et al. (2003) introduced the usage of tracking data recorded during experiments as an objective measure of presence. They reported that when agents invaded the personal space maintained by participants, the participants moved away from the agents. This is in keeping with observations by Argyle and Dean (1965). During the design phase of one of the experiments (Chapter 5), particular attention was made to obtain accurate and regular recording of time stamped tracking (positional and orientation) data from participants. In addition the position and orientation of the animated virtual humans were also recorded in order to enable a limited replay of the participant's experience.

3.4 Data analysis

Once data was collected there were three main phases to convert seemingly random information to substantial findings.

- **Coding:** The first phase of data analysing involved cleaning and rearranging the data into a convenient format. In the case of quantitative data, this was undertaken during the experiments in parallel.

- **Categorise, sort and order:** This phase involved grouping the data into categories in accordance to conditions of the experiment.

- **Deducing:** The final phase of analysing included identifying commonalities in the data, testing for significance and interpreting the results into logical conclusions.

This section is a brief overview of the types of quantitative and quantitative analytic methods that were used to uncover correlations in the collected data. Section 3.4.1 presents the statistical methods used to analyse data collected through questionnaires. Sections 3.4.2 and 3.4.3 present the methods used to analyse the physiological and behavioural data collected respectively. Section 3.4.4 discusses the method used to analyse the qualitative data collected.

3.4.1 Quantitative analysis of questionnaire responses

The most common statistical methods used in experiments designed with pre-selected information on the independent variable and continuous information on the dependent variable, are univariate analysis of variance (ANOVA), analysis of covariance (ANCOVA) or multivariate analysis of variance (MANOVA - multiple dependent measures) (Creswell, 2003). The ANOVA tests for significant differences between means. ANOVA also detect interaction effects between variables, and test more complex hypotheses about the data set. In factorial designs both interaction and main effects of ANOVA are used (Creswell, 2003). This was used to visualise the tracking data in one of the main experiments (Chapter 5). The techniques used to analyse the data collected through questionnaires in the experiments were multivariate methods since this research focused on the relationship between multiple variables. Some examples of multivariate statistical methods include logistic regression and log-linear regression (Rencher, 2002).

3.4.1.1 Coding the response variables

Coding data refers to the conversion of participant responses to a simple and consistent representation which makes the data more manipulatable by a software program such as Matlab (MathWorks, 2006) or GLIM (NAG, 2006). There were mainly two types of questionnaires utilised in the experiments: dichotomous 'yes/no' questions and ordinal Likert-type response scales.

Dichotomous response scales were the most straightforward questionnaires to code. In the dichotomous questionnaire, each of the n items on the questionnaires required the participant to respond to a 'yes/no' statement. Each item was then coded as a binary value of 1 or 0 depending on the questionnaire. The response variable is constructed by counting up the scores from each item.

There were two groups of Likert-type questionnaires used in the experiment. The questionnaires asked the participant to indicate their degree of agreement or disagreement with a statement on an ordinal[2] response scale. Scores on the response scale in both groups had a direct interpretation. High scores on the response scales corresponded to a higher response and low scores meant less. The two groups differed in the way in which the participant's response variable to the questionnaire was constructed. Most of the personality or attitude questionnaires used contained a set of n questions rated on such a scale and belonged to the first group. Each item was a statement about personal attributes. Some items on the

[2]Ordinal implies an inherent order or sequence but not in a linear sense

questionnaire were reversed in meaning from the overall direction of the scale. These are called *reversal* items. The reversal questions (i.e. questions with the word *not*) were reversed in scale to get all the questions in the same direction. The final response variable for the participant on the scale was the sum of their ratings for all of the items in the questionnaire. The total score was interpreted in a comparative manner so a participant with a response of 80 scored higher than a participant with a response of 40 but not twice as much.

The second type of Likert-type questionnaire was similar in design and context to the SUS presence questionnaires published by Slater and Steed (2000). The resulting response variable gave a self-reported presence or copresence score of the participant. The responses elicited in this questionnaire was coded in a similar manner e.g. 1-7 but this was only a means of coding. It makes no sense to add a presence response of '7' to a presence response of '4' to get a mean response, since 7 could stand for 'strongly agree' while 4 could stand for 'undecided'. In this questionnaire, the focus was on identifying participants with a high score. Similar to the personal attribute questionnaires discussed in the previous paragraph, the response variable was constructed from a set of n questions. However, each item was treated as a binary value where a high response corresponding to a 6 or 7 in the Likert-type scale was coded as a '1'. For instance, the presence response in the SUS questionnaires were constructed from a set of 5 questions. Each of the 5 questions was scored as a high or a low response based on a threshold value. In the case of the SUS, a high score was a response of a 6 or 7. The total number of high responses (r) out of a possible high responses (n) gave the response variable for the participant. In the case of the SUS, the lowest possible presence response was 0 while the highest was 5. However, in the case of reversal items, a score of 1 or 2 was coded as a high score. Once the coding of data was completed, it was vital to clean the data. Occasional errors inevitably occurred when collecting or entering data from the paper questionnaires into electronic spreadsheets. A single error in the resulting data set may seriously distort results hence it was important to sweep the entered data for errors such as invalid data or missed questionnaire responses and resolve inconsistencies. For instance, any items missed in the questionnaire were coded with a '999' in order to identify it later on in the analysis.

3.4.1.2 Analysing the questionnaire data

The ultimate goal of statistical analysis is finding and quantifying relationships between relevant variables. Variables are related if in a sample of observations, the values of those variables are correlated in a consistent manner. There are two elementary properties to define the relation between variables: *magnitude* (measures strength of a relationship) and *reliability* (measure statistical significance of the relationship). One of the techniques used to find statistical relationships between dependent and independent variables is *regression analysis*. The goal of regression analysis is to determine the values of parameters for a function that cause the function to best fit a set of data observations. Visually regression analysis is equivalent to finding the curve that best fits the data set.

Under the null hypothesis of no influence of any independent or explanatory variable on a response, the count of items in questionnaire data would have a binomial distribution with number of trials (n), and therefore a standard statistical method called *logistic regression* was used for the analysis (McCullagh and Nelder, 1989). Ordinary regression deals with finding a function that relates a continuous dependent variable to one or more independent variables. Logistic regression is a variation of ordinary regression. It is suitable to analyse dependent variables, when the observed outcome of the variable represents the occurrence or non-occurrence of an event, (usually coded as 1 or 0, respectively). It has the advantage of never using the ordinal questionnaire responses as if they were on an interval scale. Generally, the dependent (or response) variable is binary such as success/failure. The presence and co-presence response variables in the experiments may be thought of as counts of r 'successes' out of n trials, and therefore naturally have a binomial distribution[3], as required in logistic regression. In the case where the right-hand-side of the regression consists of only two factors this is equivalent to a two-way ANOVA but using the more appropriate binomial distribution rather than the Normal.

In this regression model the deviance was the appropriate goodness of fit measure, and had an approximate Chi-squared distribution with degrees of freedom depending on the number of fitted parameters. A rule-of-thumb is that if the deviance is less than twice the degrees of freedom then the model is a good fit (at the 5% significance level). The change in deviance by deleting or adding to the regression model, is particularly useful, since this indicates the significance of that variable in the fitted model. Here a large change in deviance indicates the degree of significance, i.e. the contribution of the variable to the overall fit of the regression model. The tabulated χ^2 5% value is 3.841 on 1 d.f and 5.993 on 2 d.f. If the change in deviance caused by deleting a variable from the fitted model is greater than the tabulated χ^2 5% value then the variable is significant.

3.4.2 Quantitative analysis of physiological responses

Physiological analysis were conducted by comparing the recordings of the participants' galvanic skin responses (GSR) during different sessions and conditions within the experiment. Using physiological responses as objective measures of presence in VEs is a relatively new strategy.

Experiments have shown that GSR and heart rate response could be used as objective measures in monitoring participant response in VEs (Jang et al., 2002). The study found that GSR and heart rate responses were associated with the arousal of participants in the VEs and that such measures generally returned to normal over time. Meehan et al. (2002) hypothesised that to the degree that a VE seems real, it will evoke physiological responses similar to those evoked by the corresponding real environment. Furthermore, Meehan et al. (2002) argued that higher presence will evoke a greater response, such as an increased heart rate and an increase in the participant's skin conductivity, during stressful periods, in their case when participants approached a virtual precipice. They note that heart rate and skin conductance measure the arousal of the individual, therefore might only be used when such arousal is intrinsic to the task and associate with fear.

Recent work reported by Brogni et al. (2003), Slater et al. (2003) and Guger et al. (2004) showed that physiological measures can be used for event-correlated analysis, in particular in relation to breaks in presence (BIPs). Slater et al. (2006) reported an experiment in which participants were physiologically

[3] The binomial distribution is the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p.

monitored while they experienced a bar-like environment with induced anomalies. While analysing the collected GSR data, Slater et al. (2006) looked for spikes in the data some time after the anomaly was experienced. Results indicated that GSR could be used to characterise an anomaly in a IVE. This finding was especially significant given that the VE was depicted a very busy and noisy bar scenario.

3.4.2.1 Galvanic Skin Response: coding and analysis

In this thesis, the skin conductance responses (SCR) extracted from the GSR recordings were the focus of physiological analysis. Peaks in the GSR recordings are triggered by an increase in sympathetic activity of the autonomic nervous system following the onset of an event, a sensory stimulus or a psychological stimuli (Andreassi, 1995). Momentary fluctuations in the recordings which occur with an event/stimuli are termed *phasic* responses while the relatively stable readings are referred to as the *tonic level*. The different kinds of measurements that can be extracted from GSR recordings include the level of conductance during a given period of time, the number of changes in conductance during that time, the magnitude of the response, latency (provided a time stamp of the stimuli set is maintaining) and recovery time after a stimuli (Andreassi, 1995).

The primary analysis conducted, in this thesis, compared and contrasted the peaks in GSR recordings experienced between conditions. Analysis was conducted on the participant's GSR recording by using the number of skin conductance responses (SCR) as the main variable. Since there is no standard for defining a SCR, throughout this thesis an SCR is defined as a peak rise in the GSR of at least 0.1 microsiemens over 5 seconds (Slater et al., 2006). The number of SCRs during a period of interest was calculated and analysed to uncover factor-related differences by fitting the explanatory variables to a log-linear regression model with a Poisson distribution (McCullagh and Nelder, 1989). The number of SCRs experienced by the participant depends on their particular physiology. Some people have a high rate of spontaneous SCRs (*labiles*) and others a much lower rate (*stabiles*). In order to take into account the physiology of the participant, the rate of SCRs recorded during baseline resting periods of the experiment was fitted into the regression model. Similarly, the duration of the experience was included in the model to eliminate the effect of having an increased number of SCRs due to the amount of time spent.

In addition, analysis were done with the view to detect changes in GSR with respect to events in the VE such as when a virtual human turned to look at the participant. This was done by extracting and analysing the GSR data for 10 seconds before and after the event. Each participant was exposed to 'n' number of events. The GSR data corresponding to all the events a participant experienced was summed and averaged over the 20 seconds and plotted to visualise any obvious participant responses to the event. If the event was significant to the participant, the resultant plot was expected to rise a few seconds after the event and this would be repeated over all 'n' events. In these event-related analysis, the focus was on detecting any patterns of change in the GSR data around the events.

3.4.3 Quantitative analysis of behavioural (proximal) responses

Bailenson et al. (2003) used a four-way ANOVA with absolute minimum participant-agent interpersonal distance to compute the interaction and main effects of the conditions in the experiment. A similar technique was used to compute the effects of the experiment's independent variables on the mean interpersonal distances maintained by the participant in relation to the agents in one of the main experiments reported in Chapter 5. Analysis of the proximal data was carried out using Matlab (MathWorks, 2006). Data corresponding to the periods of interest was extracted from the participant position recordings using Matlab. Initially, this was done using the event markers in the tracking data. However, a closer visual inspection of the data revealed that during interaction with agents, the participants stood facing the agents and made relatively little movements. This was indicated by a characteristic pattern in the tracking data plots as shown in Figure F.19 in Appendix F.7. These patterns were identified and extracted manually. A set of variables including the minimum and average interpersonal distances, maintained by the participants during their interactions with the agents, were computed and analysed using ANOVAs.

3.4.4 Qualitative analysis of post-experiment interviews

Qualitative methods involve making sense of textual data. There are numerous strategies developed by researchers trying to tailor qualitative data analysis. These strategies can be categorised along two dimensions: *a priori*-emerging and simple-complex (Tashakkori and Teddlie, 1998). The first of these dimensions is defined based on whether the themes used to categorise the data were established before (a prior) or during (emerging) the analysis. The second dimension defines the degree of complexity of the scheme of analysis.

The method of analysis used for the interview responses was a form of *thematic analysis* (Creswell, 2003; Tashakkori and Teddlie, 1998). The analyst defined an initial set of categories. This was done by reading through all the text before the coding process and identifying a set of broad categories which could be used to structure the data. This process also allows the analyst to note down particularly interesting observations made by the participant. Once the initial set of categories are defined, the analyst begins a detailed analysis by going through the data, categorising it, further segmenting it and labelling it with a code. ATLAS.ti (2006) was used during the coding phase of the qualitative analysis. The number of instances within the text that fitted into each category was then summarised to locate themes that related to the research questions. Since there was a set of research questions associated with the experiment, a problem-driven view of the analysis was conducted. Within this approach, the analyst starts from the research question and then looks for analytical paths from the sample data to the answers. ATLAS.ti (2006) allows the analyst to use links between codes to help define the possible relationship between categories. This facility was then used to interpret the themes and possible cause-and-effect relationships between them. These themes and relationships are then discussed in conjunction with previous findings. These steps can be generalised as follows (Creswell, 2003);

- **Organising and preparing data**: transcribing interviews, sorting and arranging the data into easily referrable groups

- **Gaining an overall perspective on the data**: reading through the data, checking for credibility, checking for depth and noting down interesting information or thoughts

- **Begin detailed analysis with a coding process**: segmenting the data into categories and labelling the categories with a term

- **Generate a description**: Using the codes to generate a small number of themes which will appear as the major findings of the qualitative studies.
- **Advance on the description and themes**: Discussing the chronology of events, detailing the themes, illustrating the themes using visuals, figures or tables.
- **Interpretation**: Comparing and contrasting the findings with information obtained from past literature to convey the meaning of the gathered data.

Video recordings and observations: Video recordings of the participant were made from the rear, throughout the experiment. The recordings showed the vocal responses and behavioural responses of the participants to the virtual humans in the experiments. The recordings were mainly made with the view to add visual detail during the analysis of the interview data. However, the recordings were ultimately used as an important source of explanatory information when irregularities were detected in the tracking data collected.

3.5 Summary

The methodology used in this thesis has been adapted to suit experiments investigating different participant responses to virtual humans portraying underlying emotional states using nonverbal behavioural cues. This chapter has focused on the experimental paradigm (Section 3.2), the apparatus (Sections 3.2.4 and 3.2.3), and the different methods that were used to collect data in experiments (Section 3.3).

Qualitative methods included questionnaires and semi-structured interviews (Section 3.3.1). Quantitative methods included physiological and behavioural measures of participant responses (Section 3.3.2). This distinction also applied to the analysis of the collected data. The interview data were analysed using a qualitative method (Section 3.4.4) while the rest of the data is analysed using statistical methods (Sections 3.4.1, 3.4.2 and 3.4.3). Traditionally these methods are used separately but in this thesis a mixed methods approach (concurrent triangulation) was taken. Generally the method calls for equal priority to be given to both types of data but in this case, priority was given to the data that were analysed using statistical methods. Qualitative results was merely used to strengthen the quantitative results or provide explanations for discrepancies. This approach fitted neatly into both the operational definition of presence and the research aim of this thesis.

A number of experiments were run prior to the main experiments described in Chapters 5 and 6. These contributory experiments include: an experiment on collaboration in across a tele-immersive environment (Mortensen et al., 2002), an experiment on social paranoia (Freeman et al., 2003) and two preliminary experiments investigating the role behavioural fidelity in virtual humans (Chapter 4). The first two experiments are not discussed in this thesis. The next chapter discusses the later two preliminary experiments. The first of these focused on the impact of gaze behaviour on the perceived quality of communication between two participants represented by avatars. The second experiment focused on the importance of perceived behaviour in agents which populate a virtual urban street. Findings from these experiments were used to improve the behaviour fidelity of agents used in the main experiments. Many of the methods described in this chapter were developed during the first two preliminary experiments.

Chapter 4

Preliminary Experiments

The experiments reported in this chapter were designed to explore the impact on participants interacting with virtual humans with varying levels of behavioural capabilities. The premise was that if subtle but significant behavioural cues are implemented in virtual humans, participants interacting with these virtual humans will experience responses which are similar to those experienced in the physical world.

Freeman et al. (2003) revealed that agents displaying minimal behavioural cues can elicit realistic responses in participants even in cases where the visual appearance of the agents are not of a high calibre. Findings in the experiments indicated that participants attributed sentience and a mental state to agents animated to display various behaviours including smiling, looking at the participant and talking to each other. Participants in the experiments typically ascribed benevolent intentions to the agents, but some participants had thoughts of a persecutory nature about the agents, although the behaviours scripted in the agents were of a neutral nature (Figure 2.1(f)). Interestingly these participants had significantly higher levels of interpersonal sensitivity (feelings of personal inadequacy and inferiority) and anxiety. This is in keeping with findings, reported by Pertaub et al. (2002), that independent of the participant's propensity to feelings of anxiety to public speaking in the physical world, participants reported feelings of anxiety to an extremely negative audience. These results suggested that participants experience realistic responses towards virtual humans.

Two preliminary experiments were designed to investigate the role of nonverbal behaviours in virtual humans through exploring participant responses to the virtual humans. The first experiment investigated the impact of implementing a model of gaze behaviour in avatars through the participant's reported social responses (Section 4.1). An additional aspect explored in this experiment was the impact of avatar photorealism. The second experiment explored the impact of varying the visual realism of various elements within a scene including agents (Section 4.2). Subsidiary aspects investigated in the second experiment included participant expectations of agent behaviour and the importance of agent responsiveness in IVEs. A discussion of the findings from both experiments is covered in Section 4.3 followed by the chapter summary in Section 4.4.

4.1 Preliminary experiment on avatar fidelity

The experiment presented in this section was designed to explore the impact of two related facets of avatar fidelity, appearance and nonverbal behaviour, on the perceived quality of communication in a shared VE. The particular modality of nonverbal behaviour focused on was gaze within dyadic interactions.

4.1.1 Hypothesis

The hypothesis was that behaviour fidelity would far outweigh the importance of visual fidelity in terms of its effects on participants' perceived quality of communication. The premise was that a simple but meaningful gaze model would improve participant responses independent of the visual appearance of the avatar.

4.1.2 Justification for using gaze behaviour

Gaze behaviour was chosen for two reasons. Firstly, the simulation of gaze behaviour in virtual humans is perhaps the most significant part of the face especially in virtual applications requiring face-to-face communication since there is evidence that a fixed stare can cause negative evaluation of the conversational partner (Argyle et al., 1974). While a simplistic random gaze model might suit the purposes of livening the virtual human, the inclusion of even a simple meaningful gaze model can significantly improve the perceived quality of communication with a virtual human (Garau et al., 2001). This made gaze the best candidate for illustrating the importance of implementing computational models of key nonverbal behaviours in virtual humans. Secondly a survey of literature revealed that gaze behaviour within dyadic situations have been studied to such an extent that viable parametric models can be reliably built (Section 2.4.2). In addition, non-parametric models have been built using data obtained through the use of an eye-tracker (Lee et al., 2002). A high-fidelity parametric gaze model, based on the parametric model used by Garau et al. (2001) and the data-driven model reported in Lee et al. (2002) under non-immersive conditions, was created for use with full-body avatars under immersive settings.

4.1.3 Experiment Design

		Behaviour fidelity	
		Random Gaze	*Inferred Gaze*
Visual realism	*Cartoonish*	6 males + 6 females	6 males + 6 females
	Texture-mapped	6 males + 6 females	6 males + 6 females

Table 4.1: The 2x2 factorial design

A between-groups 2x2 factorial designed experiment was used to investigate the impact of varying levels of avatar *visual* realism and *behavioural* fidelity on participant responses in an IVE. One genderless cartoonish avatar and two gender-specific texture-mapped avatars were used to implement two levels of visual appearance. The cartoonish avatar is depicted in Figures 4.3(a) to 4.3(c) while the texture-mapped avatars utilised in the study are depicted in Figures 4.3(d) to 4.3(i). Two levels of behaviour were simulated for the study. The first category of behaviours were designed to model realistic (inferred) gaze behaviour using the high-fidelity parametric gaze model with accompanying arm, legs and head animation. The second category of behaviours acted as the control set of behaviours and consisted of a random gaze model as well as the same arm, legs and head animations.

Forty-eight participants were recruited through poster campaigns across the UCL campus and randomly pre-assigned to gender-matched pairs. The paired participants were asked to undertake a ten-minute task together in the IVE. The same role-playing negotiation task[1] fashioned by Garau et al. (2001), was used in the experiment. Each participant was represented to the other through identical gender-matched avatars and the response explored were the participants' self-reported quality of communication at the end of the task. Garau (2003) reported complete details of the experiment design including piloting, task and procedures. The technical aspects of the experiment and its relevance to this thesis is covered in the following sections.

4.1.3.1 Apparatus

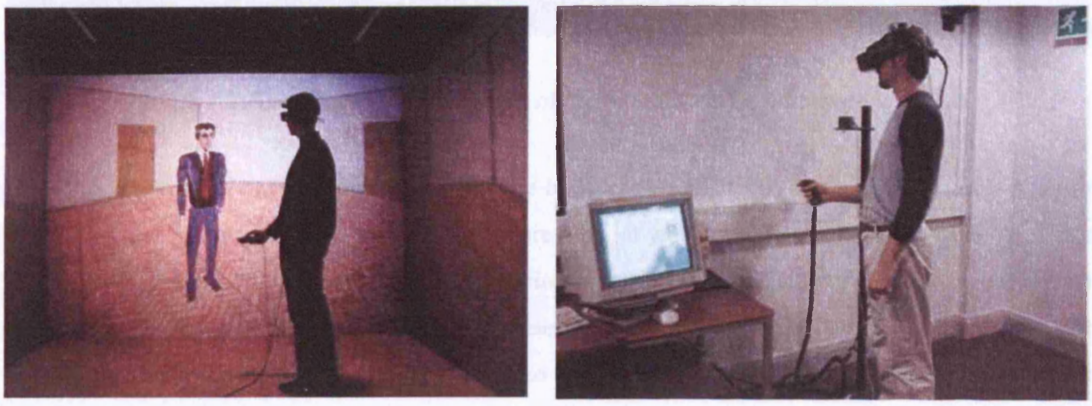


Figure 4.1: The VR systems. Left: The Trimension ReaCTor (a CAVE^{TM}-like system). Right: The Head Mounted Display (HMD).

Since there were two participants involved in each trial of the study, two types of immersive virtual reality (VR) systems were used (Figure 4.1). One participant operated in a ReaCTor while the other used a head-mounted display (HMD) (Sections 3.2.3.1 and 3.2.3.2). As can be deduced from the images (Figure 4.1), the participant using the ReaCTor did not see their virtual self while the participant using the HMD did not see their real body. In addition, the rate at which data from the trackers was used to updated the virtual environment was kept constant at 10Hz for both the ReaCTor and the HMD. This was a further precaution to ensure results obtained were purely due to the implementation of the gaze behaviour in the avatars and not a difference in the quality of display systems or fluidity of the body animations.

Participants were also given a wireless microphone in order to facilitate an audio link (verbal conversation) with each other in the IVE. Existing software was used to distinguish between speech and background noise in the audio data from the microphones. The software was calibrated by setting a threshold value for each microphone at the beginning of the day. Audio input higher than the set thresh-

[1] Participants in the dyad were randomly assigned the role of a mayor or baker and asked to settle a potential family scandal.

old indicated the presence of speech. The presence of speech was used to control gaze behaviour by setting the internal state of the avatars to either *talking* or *listening*. These states are discussed in more detail in Section 4.1.5.

4.1.3.2 Software

The virtual humans and environments were implemented on DIVE (Section 3.2.4.1). Two C++ DIVE plugins were used extensively in the creation of animations for all the virtual humans used in the experiment. The algorithms used are discussed in Section 4.1.5.

4.1.3.3 The Virtual Environment

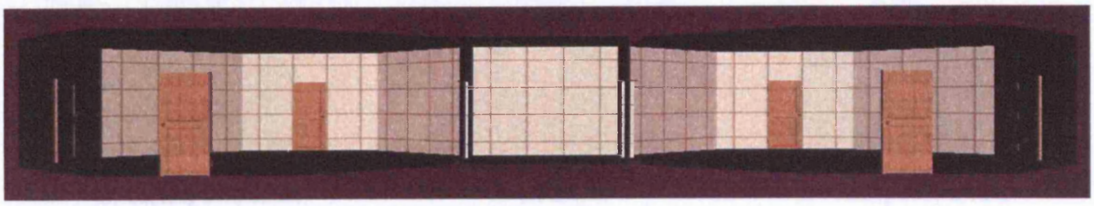
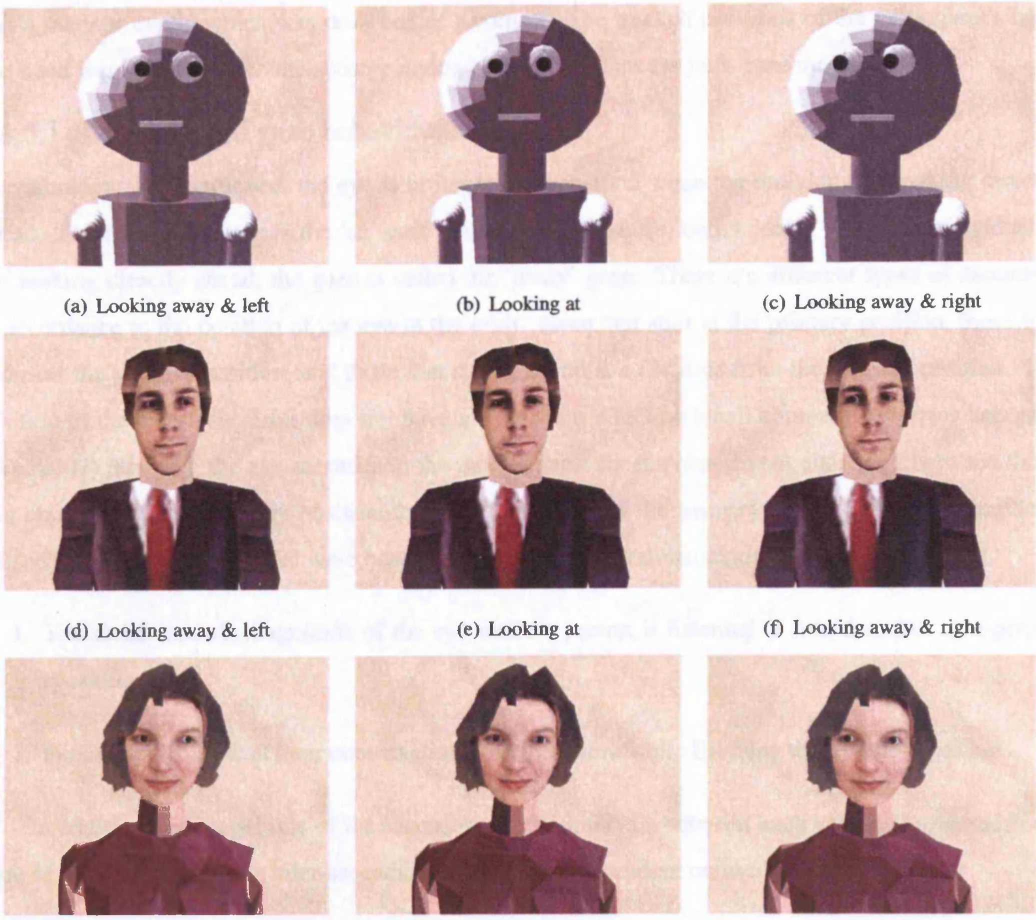


Figure 4.2: Longitudinal cross section of the virtual training and meeting rooms

The experiment was conducted in a purpose-built low-detailed VE made of two training rooms connected to a meeting room in the middle (Figure 4.2). It was important to have a sparse VE, since in the physical world, the more features in an environment the more visually distracted the participant is likely to get, hence diverging from the parameters described for gaze behaviour in literature. The training rooms, at the far ends of the VE, were used to familiarise each participant with navigating their VR system. In this way, participants were kept separated until the start of the experiment. At the start of the experiments, virtual doors, separating the training rooms from the central meeting room, were opened allowing participants to make visual contact with each other's avatar and interact. This was facilitated through the DIVE-Tcl interface mentioned in Section 3.2.4.1

4.1.4 The Avatars: Visual appearance

Three H-Anim (2006) compliant avatars were designed for use in this study, one of which was visually very cartoonish and gender-neutral (Figure 4.3(a) - 4.3(c)). The cartoonish avatar was built based on mathematical proportions used by artists (Larmann, 2006). The dimensions of all the body parts were described as a multiple of the size of its head. For instance, the height of the avatar was eight times that of its head. The whole avatar could be scaled to a required height by manipulating one variable (Appendix I.3). This was useful in making sure that the cartoonish avatar was of the same size as the other two texture-mapped gender-specific avatars (Figures 4.3(d) - 4.3(i)). An additional aspect taken into account, in the design of the texture-mapped avatars, was the subtleness of features on the face in order to avoid attracting attention to any one feature. In contrast, the eyes on the cartoonish avatar were emphasised in order to encourage the participant's attention to the eyes and less on the lack of other facial features. It was vital to avoid other accompanying facial expressions, including lip synching, to ensure any resulting effects were purely due to the gaze behaviour model. The texture-mapped virtual



(a) Looking away & left (b) Looking at (c) Looking away & right

(d) Looking away & left (e) Looking at (f) Looking away & right

(g) Looking away & left (h) Looking at (i) Looking away & right

Figure 4.3: Gaze behaviour in the three avatars (genderless cartoonish, male texture-mapped and female texture-mapped): looking left, onward and right

humans used in all the experiments presented throughout this thesis were adapted from Pertaub et al.'s (2001) study on the fear of public speaking in virtual environments.

4.1.5 The Avatars: Behaviour generation

The behaviour module was designed to autonomously control important behaviours including those that are not consciously driven such as gaze behaviour. It was designed to adapt the behaviour of the avatar depending on the status of the participant based on real-time speech and tracking data. The participants were not aware that they are controlling the module indirectly. The module consisted of an eye-gaze model and other simple body animations.

There were two algorithms implemented within the gaze model. Only one algorithm was used per experiment condition. The first, an inferred model, was designed to provide a realistic gaze simulation that mimicked the gaze behaviour of individuals in a real dyadic communication. In this module, certain internal states of the avatar were used to control the realistic gaze behaviour in the parametric model at run-time. The behaviour module continuously checked these states, to determine (a) if the avatar was talking or listening and (b) if the avatar was looking directly ahead (primary position) or not. The second was a random model which was designed as a control. The tracked positions of the participant's head and hand were used to drive the accompanying body animations for both gaze models.

4.1.5.1 The inferred gaze behaviour model

Terminology: As mentioned, the eye is in its primary position when the individual is looking directly ahead. The associated gaze is the '*at*' gaze (Figures 4.3(b), 4.3(e) and 4.3(h)). When the individual is not looking directly ahead, the gaze is called the '*away*' gaze. There are different types of saccades[2] in accordance to the position of the eye in the orbit: those that start at the primary position, those that end near the primary position, and those that start and end at a distance from the primary position. The position of the eye in the orbit does not have a significant effect on small naturally occurring saccades (Lee, 2002) therefore the eye saccades in the module built for the experiment alternated between those that started from the primary position to those that ended at the primary position. The assumptions behind the inferred gaze model were based primarily on observations reported by Argyle (1998):

1. The mean saccade magnitude of the eye while a person is listening is less than that of a person speaking.

2. People tend to look at their conversational partner more while listening than while speaking.

In addition, the magnitude of the saccades and the intervals between each saccade varied with the state of the participant. The inter-saccadic intervals were dependent on two factors:

1. Whether the participant was speaking or listening and

2. Whether the avatar's eye was in the primary position or not

In the behaviour module, with the advantage of tracking in the IVE, participants were left in full control of their avatar's head. The gaze behaviour was kept independent of the head movements unlike Garau et al. (2001) and Lee et al. (2002). Therefore even when the gaze behaviour model determined the avatar's eye to be in the primary position, looking directly ahead, the avatars were in mutual gaze only if each participant chose to look at the avatar of their partner and vice versa. Figure 4.4 depicts the algorithm used to animate the inferred gaze model.

In the inferred gaze model, the eye began at the primary position at the start of the conversation. During the conversation, each avatar's behaviour module checked status-monitoring variables to determine if the avatar should be in '*talking*' mode or in '*listening*' mode. It also checked if the avatar should be in '*looking at*' or '*looking away*' mode. Once the status of the avatar was determined, the appropriate inter-saccadic interval was calculated on the basis of an exponential probability distribution, about the mean times given in Table 4.2. The position of the eye was paused for the duration of the resultant inter-saccadic interval. At the end of the inter-saccadic interval, the magnitude and the direction of the next saccade, if not to the primary position, was calculated.

[2]Section 2.4.2 discusses saccades and related parameters such as inter-saccade intervals, saccade duration, saccade magnitude, saccade direction and saccade velocity.

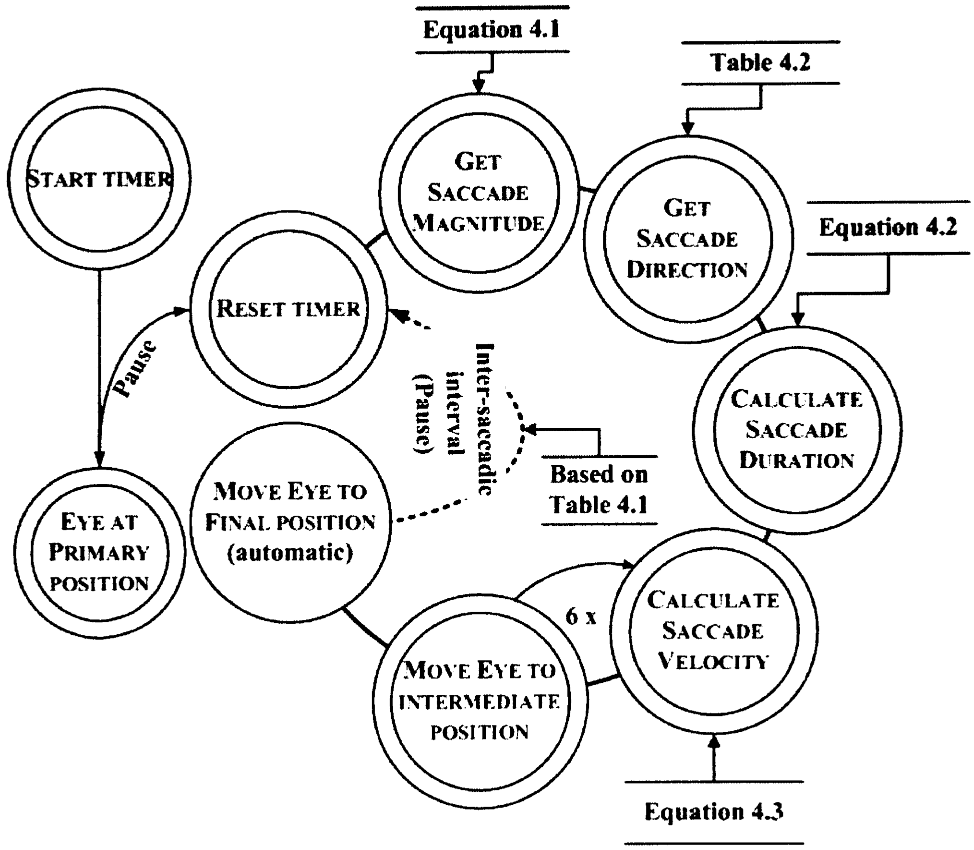


Figure 4.4: Overview of the inferred gaze model.

	Looking At (s)	Looking Away (s)
Listening	2.5	1.6
Speaking	1.8	2.1

Table 4.2: Mean inter-saccadic intervals corresponding to the state of the avatar

According to Lee et al.'s (2002) data-driven model, the frequency[3] of a particular saccade magnitude occurring can be fitted onto an exponential function. The saccade magnitude of an *away* gaze was calculated as shown in Equation 4.1 where the variable '*Rand*' was a random number from a uniform distribution between 0 and 15 (Lee et al., 2002).

$$Magnitude = -6.9 * ln(Rand/15.7) \tag{4.1}$$

The maximum magnitude of a saccade in the distribution was dependent on whether the participant was talking or listening. While a person was talking the maximum saccade magnitude was set at 27.5°. This reduced to 22.7° when the participant was listening. These parameters are consistent with Bahill et al.'s (1975) theory that 90% of all natural saccades are less than 15°. The duration of each saccade was dependent on the magnitude of the saccade (Lee et al., 2002) as given in Equation 4.2. Experimental

[3]Unlike Lee et al. (2002), all timing functions within the behaviour module were implemented in terms of absolute times (precision to microseconds) rather than frame rates. This ensured that the timers would not be influenced by varying refresh rates.

evidence suggested that the value of Δ in Equation 4.2 is between 20 - 30 milliseconds and Lee et al. (2002) chose a constant value of 25 milliseconds. The variable δ in Equation 4.2 is the incremental slope.

$$Duration = \Delta + \delta * Magnitude \tag{4.2}$$

$$\begin{aligned} \Delta &= 25ms \\ \delta &= 2.4ms/degree \end{aligned}$$

The accompanying saccade direction of an *away* gaze was chosen out of eight distinct directions evenly spaced at 45°. Provided the next saccade was not to the primary position, its direction was determined using a uniform random number generator and a list of probabilities based on Lee et al.'s (2002) findings (Table 4.3). The saccade direction was independent of head rotation.

Direction: anti-clockwise from right	0°	45°	90°	135°	180°	225°	270°	315°
Percentage of occurrence (%)	15.5	6.5	17.7	7.4	16.8	7.89	20.4	7.8

Table 4.3: Occurrence of saccade directions in the inferred gaze model

The final variable in the inferred gaze model was the velocity of the eyes during a saccade. In reality, an individual's eyes do not maintain a constant velocity. They accelerate from the start position to a maximum velocity and then decelerate to the end position of the saccade. A higher magnitude results in a higher average saccade velocity. Lee et al. (2002) constructed a polynomial function to calculate the instantaneous velocity of the eye during saccades, however, the function did not yield the expected numerical results. Therefore an alternative model was designed in which instantaneous velocities were calculated along the path of each saccade after it was divided into six equally spaced frames (Equation 4.3).

$$Y = 14 + exp\left\{-\frac{\Pi}{4} * (X-3)^2\right\} \tag{4.3}$$

$$\begin{aligned} Y &= \text{Instantaneous Velocity at Frame X} \\ X &= [1, \ldots, 6] \end{aligned}$$

The six frames in the saccade were set based on the saccade duration calculated in Equation 4.2. The eye was then moved to the corresponding intermediate positions at the end of each frame depending on the elapsed time and the instantaneous velocity. Figure 4.5 depicts the instantaneous saccade velocities predicted by the alternative model, which closely mimics the data reported by Lee et al. (2002). Since the intermediate shifts in eye gaze was normalised to the saccade magnitude of the eye at the time, the eye followed the curve and moved to the corresponding saccade position. When the eyes reached the

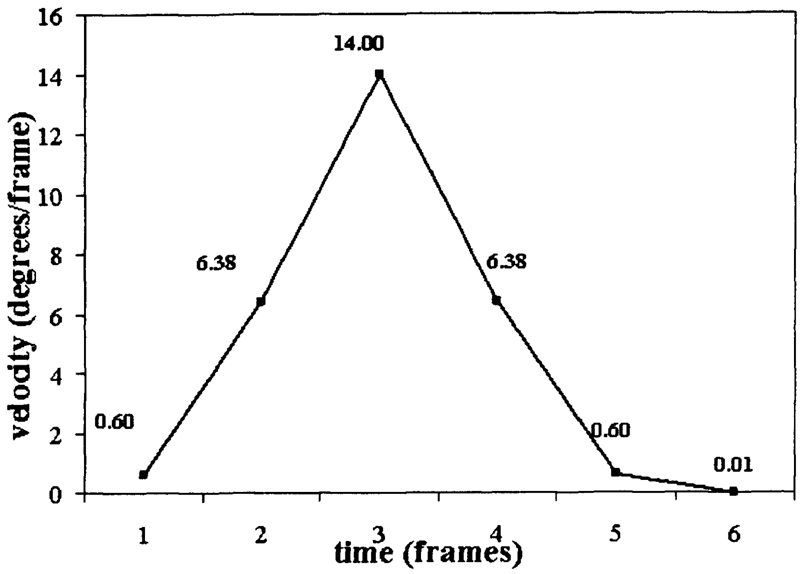


Figure 4.5: Instantaneous saccade velocities produced using the alternative model.

end of the saccade, the whole process of finding a new saccade magnitude, direction, duration and an inter-saccadic interval was repeated as depicted in Figure 4.4.

4.1.5.2 The control - a random Gaze behaviour model

A random gaze model was implemented for use in the control conditions of the experiment. The main difference between the random and inferred gaze models were that the random gaze model was not controlled using the participant's speech pattern. The inter-saccadic interval was a constant 2 seconds independent of whether the avatar was talking or listening, looking at or looking away from the conversational partner. In other words, the random gaze model was not tied to conversation between the participants. The inter-saccadic interval (2 seconds) was the mean of all the inter-saccadic averages from the inferred eye-gaze model. This ensured that the flow of the gaze animations was not too fast or too slow in comparison to the inferred gaze model.

The model was based on the same variables as the inferred gaze model except there were no saccade velocities involved. The magnitude of the saccade was obtained using a uniformly distributed random number generator. In order to avoid having a model that was too unrealistic, the saccade magnitude was limited to a maximum of 15^o. The direction of the saccade was implemented as a random direction uniformly distributed around a 360^o circle centred at the primary position as opposed to the 8 distinct directions in the inferred gaze model. The differences between the resultant animations between the inferred and control gaze models were subtle and not easily noticeable to the casual observer.

4.1.5.3 Accompanying body animations

Some simple motion algorithms were incorporated into the behaviour module with a view to preserving visual consistency between the movement of the eyes and the rest of the avatar's body. Animations were created for the avatar's head, right arm and legs. The complexity of the body animation in the avatars was limited since the only body parts tracked in both IVE systems were the head and the right hand. The body animations of the avatars semi-mirrored the motions of the participants in the shared VE depending on the information obtained from the limited number of tracking sensors used on the participant (head and right hand). This is similar to the concept presented by Badler et al. (1993).

However, there was an advantage in the VEs to counterpart the lack of tracking. The participant using the systems did not see both their real body and their animated virtual self at the same time. This meant that when animating the avatar, more importance could be placed on visual consistency over accuracy in mirroring the body of the participant.

Head Control: The direction of travel in the virtual environments was based on the direction of movement of the participant's head tracker in the X-Z plane. Initially, the avatar simply imitated the rotation and translation of the participant's head tracker in the X-Z plane. However, it was noticed during pilot studies that the head trackers in both VR systems were not placed centred above the participant's head. This gave the impression that a translation was associated with every head nod and head shake gesture made by the participants. Additionally the body of the avatar was being moved along with its head during these conversational gestures since the movement was interpreted as a combination of a translation and rotations instead of only rotations. The avatar's body was animated by rotating from side to side during a head-shake gesture e.g. when the participant gestured 'no'. If the participant nodded their head, e.g. to gesture 'yes', the whole avatar moved forward and backward rapidly.

Simple conversational gestures such as head nods and shakes play an important factor in social interactions. Therefore a simple rule in the behaviour module was used to define a boundary circle around the position of the avatar's head. The boundary circle had a radius equivalent to seven centimetres in the X-Z plane. This value was chosen after testing both VR systems to determine the optimal threshold that could be used to differentiate between a participant attempting to gesture and a participant attempting to physically move. When the head tracker moved a distance greater than the radius of the boundary circle ($>$ 7cm), the avatar rotated to the direction of travel of the participant and mirrored the distances moved by the head tracker, hence representing the position of the participant in the shared VE. However, when the motion detected by the participant's head tracker was smaller than the boundary radius ($<$ 7cm), the behaviour module prevented the body of the avatar from translating along the ground but allowed the avatar's head to correctly emulate the rotations of the participant's head. This allowed the avatar to accurately represent head gestures using only the rotational data from the participant's head tracker. Participants could take full advantage of any head gestures, they would have normally used in the physical world, in the virtual world.

Arm and Leg movements: In addition to the head animations, the right arm and legs of the avatars were animated to follow the movements of the participant. In order for the behaviour module to accurately animate the avatars limbs, the correct height of the avatar and the length of the avatar's limbs were deduced in the first virtual frame and initialised. This allowed the behaviour module to deduce if the participant was bending their right arm or crouching in latter frames. It also allowed for deductions of simple situations such as the participant taking off their trackers or positional data from one of the trackers not being received correctly. The right arm was animated based on a deduction of joint locations using the positions of the hand and head of the avatar.

The position of the avatar's shoulder was always known due to the hierarchical nature of the avatar. In accordance to the hierarchical structure of the H-Anim (2006) complaint avatars, the shoulders of the avatar were children joints of the body. Since the body followed the X-Z position of the head tracker, it was possible to deduce the current position of the right shoulder using an object location functionality in DIVE (SICS, 2006). The position of the right wrist was always obtainable due to the data collected from the right hand tracker embedded in the navigational joystick in our systems. The vector position of the right elbow of the avatar was then deduced using the 3D vector co-ordinates of the right shoulder and the right wrist of the avatar using an algorithm similar to inverse kinematics. The 3D position of the right elbow of the avatar, projected down onto the axis between the right shoulder and the right wrist, was deduced using the cosine rule. In addition, some geometrical deductions were made to ensure the tracked right wrist of the avatar was not detached from the body of the avatar e.g. when the participant dropped the VR navigation wand. The deduced vector co-ordinates of the projected elbow was then cast onto a direction orthogonal to that of the axis defined by the right shoulder and the tracked wrist. Once the right elbow position of the avatar was deduced, the geometries of the right upper and lower arm were aligned between the appropriate joints to reproduce a possible 3D posture of the avatar's right arm in each virtual frame. There was a special consideration when the arm of the participant was longer than that of the corresponding avatar. In this case, the directional data from the tracker was used to reproduce the correct posture of the right arm but the translation of the right hand tracker was not conveyed to the right wrist of the avatar. During pilot studies, this methodology for manipulating the right arm was found to be adequate.

As mentioned, the behaviour module inferred information about the dimensions of the avatar as part of the initialisation process. The legs of the avatar were animated using deductions based on the height of the participant. Once a reduction in height of the participant is detected, the behaviour module used the translation data from the participant's head tracker to lower the position of the body of the avatar (including the head) in all axes. The current position of the avatar's hips during the virtual frame was obtained. As the 3D coordinates of the hips were determined in each frame, the positions of the knees were bent in the direction of the Z-axis based on total leg length. The feet of the avatar however, are only updated in the X-Z plane i.e. the Y co-ordinates of the feet are reset to null. The upper and lower leg geometries were aligned in between the appropriate joints of the avatar. All body animations were updated for every frame. This gave the illusion of the avatar crouching whenever the participant bowed or crouched in the physical world.

4.1.6 Data collected and Results

A number of subjective indicators were collected at the end of the experiment through 7-point Likert-type close-ended questions where a score of '1' corresponded to strong disagreement and '7' corresponded to strong agreement. The main indicators measured that participant's perceived quality of communication through four criteria:

- *Face-to-face*: The extent to which the conversation was like a real face-to-face interaction.
- *Involvement*: The extent to which the participants experienced involvement in the task.
- *Copresence*: The level of copresence experience - that is, the sense of being with another person.
- *Partner Evaluation*: The extent to which the participants positively evaluated their partner.

The other indicators included the participant's subjective evaluations of the avatar's gaze behaviour, overall behaviour fidelity, visual appearance, and presence. A number of control variables were also collected including the participant's age, gender, occupation, experience with computer games, experience with VR systems and their degree of participant's social anxiety as measured by the SAD questionnaire (Watson and Friend, 1969, 1987). The logistic regression method described in Sections 3.4.1.1 and 3.4.1.2 was used to analysis the data collected. The analysis demonstrated a very strong interaction effect between the type of avatar and the type of gaze behaviour used. As expected, in the more realistic texture-mapped avatar, the inferred gaze behaviour increased effectiveness. However, in the cartoonish avatar, the inferred gaze behaviour *reduced* reported face-to-face effectiveness. A detailed analysis of the data collected in this experiment was reported by Garau (2003).

Garau's (2003) results suggested that there must be consistency between the visual appearance of the avatar and the type of behaviour that it exhibits. Additionally the analysis revealed that the greater the degree of social anxiety, the lower the face-to-face effectiveness. The age of the participant (Age) was also found to be significant, and positively associated with participant evaluations older people were more likely to have rated their experience as like face-to-face. The relationship between face-to-face effectiveness and the independent variables 'type of avatar' and 'type of gaze' were the same whether or not these additional explanatory variables (Age, SAD) were included.

In addition to the relationship between appearance-behaviour avatar fidelity revealed in the questionnaire data, participant's feedback on their experience during the study suggested that the avatars were able to induce realistic responses despite the relatively low level behaviour of the avatars:

- *I was surprised at how much I felt accountable to my partner's avatar, I was unable to fully take the aggressive alternative '3' as planned*
- *I wonder whether the other person could realise that I could not stop smiling at his ridiculous claims and comments!*
- *Towards the end of the conversation my partner was looking around the room and the ceiling which gave me the impression she was not concentrating, the response I had to this was very realistic, as if a real person was not concentrating on what I had to say.*
- *Once I saw my partner entering the room I became focused on him and his voice, the rest of the room I was aware of but was concentrating on my partners movement.*
- *Having anything in the environment - even if it is not a very realistic avatar or room - helped a little. It gives something to focus on. Although you do not think of it as a person, strangely it does*

stop you turning away or doing anything inappropriate. Also your mind does not wander so much as it might on the telephone. You are immersed in the environment even if it never becomes reality.

However, a lack of expression in the avatars did have a negative impact on participant responses:

- *A large part of a normal conversation - especially given the delicate nature of the subject, involves a lot of facial expressions & gestures, which play a large part in the conversation. With the absence of the two, voice and the subtle changes in tone became a much more important part of conveying feeling and mood. After realising the fact that the avatar conveyed very little of the persons actual physical/emotional state, it became even less believable, and I found myself concentrating more and more on the voice. Also, the slightly jerky display distracted from the sense of reality and "being there" - especially as all movements were completely virtual.*

- *... the avatar did not seem to appear life-like at all; and his movements did not seem natural; the face was expressionless.*

- *My partner's avatar did not appear to gesticulate at all with his hands or shift standing position on the floor which made it seem occasionally inexpressive and served to remind me that I was in an experiment (the 'laboratory').*

4.1.7 Summary

The experiment, presented in Section 4.1, was designed to explore the impact of two aspects of avatar fidelity on participant's self-report quality of communication within an immersive shared VE.

The first property of avatar fidelity investigated was its visual appearance (cartoonish versus texture-mapped). The second property was the avatar's nonverbal behaviour fidelity with respect to gaze. Two levels of gaze behaviour were implemented: an inferred and a random gaze model. Findings from the experiments suggested that an inferred gaze model can significantly improve the participant's perceived quality of communication. In addition the results are in keeping with previous findings by Nowak and Biocca (2003), Tromp et al. (1998) and Slater and Steed (2002) which suggest that an avatar with higher visual complexity demands a more realistic behaviour model. Finally, participant comments suggest that even though these relatively simple avatars are capable of eliciting realistic responses in participants, a lack of expression can have a negative impact on participant responses.

The experiment presented in the next section was designed primarily to explore the extent to which visual consistency within a scene is important. However, a secondary goal in the experiment, sought to investigate the impact of having no inferred behaviours in agents, on participant responses, through the analysis of semi-structured post-experiential interviews.

4.2 Preliminary experiment on visual realism

It has been argued that even simulations with a lower degree of realism can contain important cues necessary to give a believable experience such as those perceived in flight simulators or in stressful scenarios (Meehan et al., 2002; Pertaub et al., 2002; Zimmons and Panter, 2003). The experiment presented in this section focused on trying to investigate the impact of different texture quality and the visual quality of agents on the overall level of presence in a populated urban IVE.

4.2.1 Hypothesis

The experiment focuses on the perception of visual fidelity in urban VEs. The main goal was to investigate whether there needs to be consistency between the levels of realism of the different elements within a scene. In this case, the primary research question was whether the level of realism of the buildings needed to be consistent with the level of visual realism of the agents populating the environment. The hypothesis was that less repetitive textures in the scene and more visually realistic agents would enhance the participants' sense of presence. The visual realism of the IVE was varied by altering the number of textures used on the buildings and by using two types of agents. Although both types of agents were not highly realistic in terms of their appearance, one was deliberately designed to be cartoonish while the other type was texture-mapped. For instance, the second kind had a face that was texture-mapped from real human faces.

A subsidiary goal was to investigate the impact of interacting with agents with limited expression as measured through semi-structured participant interviews. The experiment was also used to test an agent animation library - Platform Independent Architecture for Virtual Characters and Avatars (PIAVCA - Section 3.2.4.3) in conjunction with DIVE (Gillies et al., 2005). Finally the experiment was used as a means to explore the use of physiological devices to gather autonomic participant responses to virtual humans.

4.2.2 Experiment Design

The experiment was designed to assess the limits of visual realism in enhancing the believability of VEs. A gender-balanced between-group two-by-two factorial design was used to explore two aspects of the VE:

- The number of textures used in the VE
- The visual realism of the inhabiting agents

		Virtual Environment Textures	
		Repetitive	*Non-Repetitive*
Agent Visual fidelity	*Cartoonish*	5 males + 5 females	5 males + 5 females
	Texture-mapped	5 males + 5 females	5 males + 5 females

Table 4.4: The 2x2 factorial design

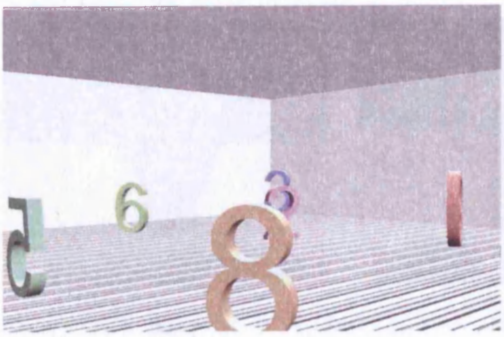
4.2.2.1 Apparatus

The Trimension ReaCTor described in Section 3.2.3.1 was used to generate the IVE. However, in addition to the mini-joystick used for navigating the VE, one of the five buttons was enabled. The usage of the active button is discussed in Section 4.2.2.6. In addition to the ReaCTor, the ProComp+ device was used to measure participants' physiological responses.

4.2.2.2 Software

The VE and agents were implement in DIVE (Section 3.2.4.1) while a VRPN client and server software (Section 3.2.4.2) were used to record participants physiological responses during the experiment. The agents were animated using an early version of Gillies et al.'s (2005) PIAVCA (Section 3.2.4.3). A single cyclic walking motion was used to animate the agents.

4.2.2.3 The Virtual Environment



(a) The Training Room

(b) An Empty Virtual Street

Figure 4.6: The Virtual Street

A virtual world was created in 3D Studio Max to resemble a typical street-like environment. The environment consisted of a high street lined with buildings on either side with a few secondary streets off the main street. The ends of the main and the secondary streets were sealed off in order to encourage participants to stay within the limits of the designed world. A few tall buildings were modelled over the horizon to maintain the illusion of a curtained off high street.

One of the factors in the experiment was dependent on the number of different textures used in the scene. In the conditions with 'Repetitive Textures', 20 textures were used to provide shop fronts whereas in the conditions involving 'Non-Repetitive Textures' there was twice the number of textures (~40). The same street model was used throughout the experiment with the only difference being the number of different textures and the type of agents used.

4.2.2.4 The Avatars: Visual appearance

There were two types of agents used depending on experiment conditions. The first type were cartoonish (Figures 4.7(a) - 4.7(c)) and the second were texture-mapped (Figures 4.7(d) - 4.7(f)). In each trial of the experiment, all the agents were H-Anim (2006) complaint and were of similar visual complexity. The street was populated with a total of sixteen agents, however, only eight agents were on the street at any one time. The decision to activate only eight agents at the same time was made due to a technical constraint in rendering all the agents in run-time at an acceptable frame rate.

4.2.2.5 The Avatars: Behaviour generation

All the agents in the experiment were programmed to behave in the same way. The agents were programmed to walk out onto the street through one of four doors at the ends of the main street. The agents

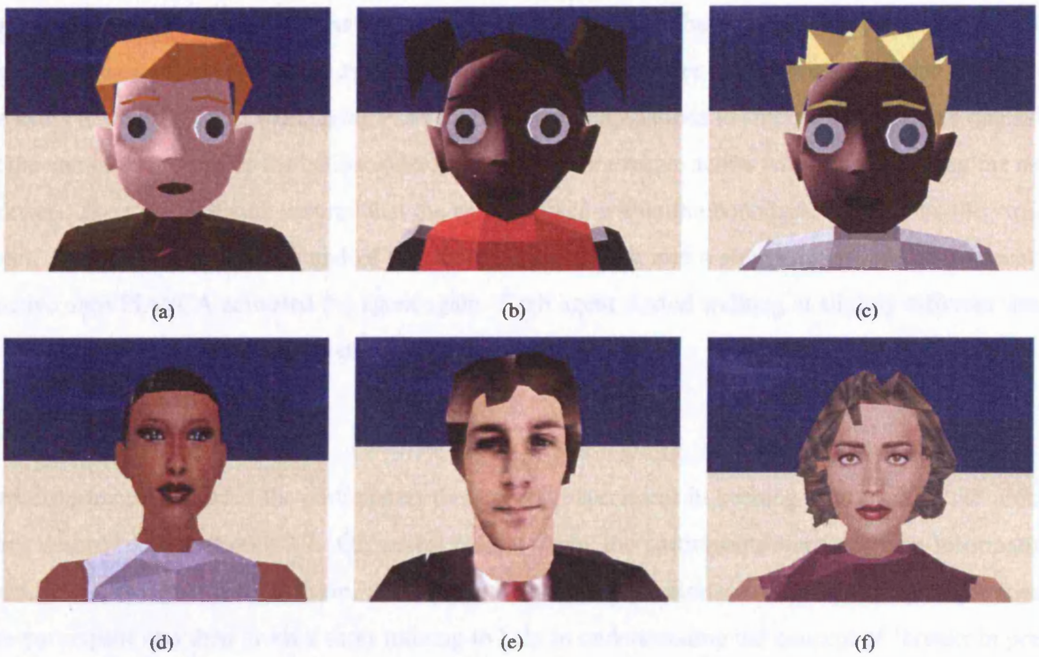


Figure 4.7: A sample of the agents: Cartoonish (a - c) and Texture-mapped (d - f)

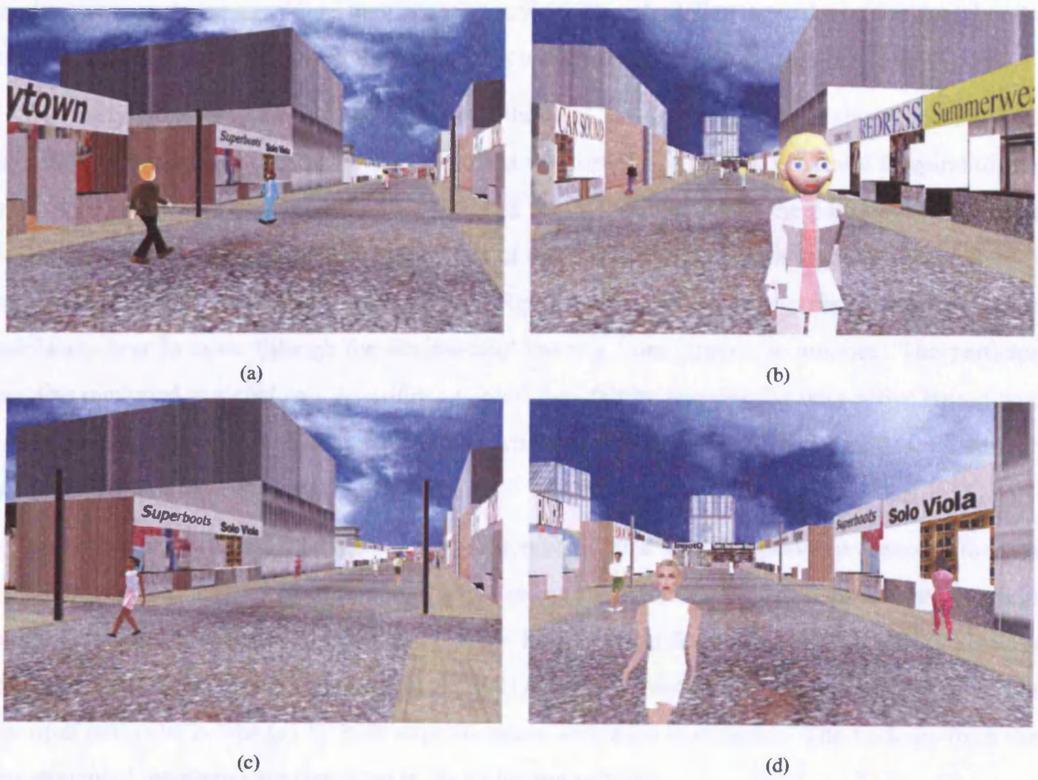


Figure 4.8: Agents on the Virtual Street: Cartoonish (a and b), Texture-mapped (c and d)

walked to the opposite end of the street using a cyclic walk pattern defined by a motion file taking care to avoid each other and the participant. A callback within the PIAVCA system (Section 3.2.4.3) was used to animate the agents walking down the street and avoiding collisions with each other and the participant.

Each of the eight active agents was aware of the position of the other agents. The participant's avatar was treated as another agent in the system. If any of the other agent or the participant were in the path of the agent it would alter its path, using PIAVCA motion manipulations to smoothly sidestep to one side. At the end of each footstep the callback determined whether evasive action was required during the next footstep. The same callback ensured that the agent walked within the boundaries defined by the virtual street. If an agent reached the end of the street it would walk into a shop entrance and then become inactive until PIAVCA activated the agent again. Each agent started walking at slightly different times so that the animation of the agents did not appear synchronised.

4.2.2.6 The Procedure

Two experimenters guided the participants through the experiment in keeping with most of the procedures described in Section 3.2.7. On arrival for the study, the participants were given an information sheet, asked to sign a consent form, and complete an online questionnaire for demographic purposes. The participant was then given a short training to help in understanding the concept of 'breaks in presence (BIPs)' (Brogni et al., 2003). In this training they were asked to look at four Gestalt pictures and switch their focus from one perceived image to the other. The transitions they experienced in the viewing exercise were equated to transitions they might experience from the VE to the real world of the laboratory during the experiment. Through out the study, BIPs were referred to as '*transitions to real*'.

The participant was then invited to step into the ReaCTor and the experimenters fitted the participants with physiological sensing devices to collect the participant's ECG, GSR and Respiration measures. A physiological baseline was then recorded for $1\frac{1}{2}$ minutes while the participant stood in the ReaCTor with the lights switched off. At the end of the baseline period, a training room containing 3D numbers appeared on the walls of the ReaCTor (Figure 4.6(a)). One of the experimenters showed the participants how to move through the environment moving from number to number. The participant were also reminded to signal any '*transitions to real*' they felt by pressing the only active button on the navigation wand. At the end of the training, the participant was told to enter the virtual street through a door and explore it for 3 minutes (Figure 4.8).

The participant was asked an immediate question about their overall reported presence in the street. The participant was then asked to complete an online questionnaire designed to gather data on various matters including their reported sense of presence based on the SUS (Slater et al., 1998; Slater and Steed, 2000) and the ITC-SOPI (Lessiter et al., 2001). The session concluded with an audio taped semi-structured interview conducted by both experimenters with each participant. The findings from these semi-structured interviews are presented in the following sections.

4.2.3 Data collected

A mixture of questionnaires, interviews, physiological measures, and BIPs were used to collect participant responses. However, analysis of the physiological measures are not reported in this thesis since they were administered by a colleague outside the scope of this thesis.

4.2.3.1 Subjective

A variety of questionnaires were administered to assess the behaviour and views of the participants:

- *Demographics:* age, gender, occupation, language proficiency, experience in computer games, programming and virtual reality etc.
- *Presence:* eleven questions from the Slater-Usoh-Steed (SUS) questionnaire (Slater et al., 1998; Slater and Steed, 2000) and the ITC Sense of Presence Inventory (ITC-SOPI). (Lessiter et al., 2001).
- *Others:* participant's perceived realism of the environment, perceived visual realism and behavioural fidelity of agents, perceived expressiveness of the agents and level of reported interaction with the agents.

The Presence-related response variables were of particular interest and were collected using an online questionnaire. Reported presence was collected using two questionnaires: SUS and ITC-SOPI. The SUS score was the number of 'high responses' out of 5 questions (Slater and Steed, 2000) while the ITC-SOPI was measured through the count of high responses to 6 questions (Lessiter et al., 2001). The reported presence responses were analysed using the logistic regression method described in Sections 3.4.1.1 and 3.4.1.2.

4.2.3.2 Objective: Physiological recordings

Physiological measures were recorded during the participant's experience, using non-invasive devices (Sections 3.3.2.1 and 3.2.3.3), however, analysis of the recordings, with respect to this experiment, are not discussed in this thesis.

4.2.4 Basic findings

The independent variables, in this experiment, were the variety of textures (repetitive and non-repetitive) and the visual appearance of agents (cartoonish and texture-mapped). The explanatory variables collected included a range of participant personal data and physiological measures.

	Cartoonish	**Texture-mapped**
Repetitive Textures	2.2±2.04	0.4±0.70
Non-Repetitive Textures	2.3±2.26	2.3±1.77

Table 4.5: Mean of SUS Presence response: maximum score = 5

	Cartoonish	**Texture-mapped**
Repetitive Textures	3.9±2.28	2.4±0.97
Non-Repetitive Textures	3.4±2.27	3.6±1.71

Table 4.6: Mean of ITC-SOPI Presence response: maximum score = 6

Tables 4.5 and 4.6 give the mean presence scores using the high responses count method as collected using the SUS and the ITC-SOPI questionnaires respectively. The questionnaire data collected was analysed using logistic regression (Sections 3.4.1.1 and 3.4.1.2). The analysis revealed that the lowest reported presence occurred in the conditions with repetitive textures and texture-mapped agents. This seems to suggest that visual consistency within a scene is important when the visual fidelity of the agents is relatively higher. Most interestingly, analysis of the questionnaire data revealed that the perceived behavioural realism of the agents was highly significant and positively associated with participant's reported presence. In other words, the more that the agent's behaviour was perceived as being real, the higher the participant's reported presence. This was the most significant factor of all, contributing the greatest to the overall fit of the regression model (deviance (χ^2) = 27.12 on 1 d.f.). It is important to note that the agents' behaviour was the same for all the conditions in the study. A further analysis of the interview data collected in the experiment with particular focus on perceived behaviour and responses is reported in the following section.

4.2.5 Findings from Interviews

Even though 40 participants took part in the experiment, 6 were excluded from the analysis since the recordings of their interviews were of poor audio quality making them unsuitable for transcription. The resulting 34 interviews were distributed across the conditions as shown in Table 4.7. Each interview was transcribed verbatim and then analysed in the manner described in Section 3.4.4 in order to explore the effects of the agents on participant responses.

	Cartoonish	**Texture-mapped**
Repetitive Textures	5 males + 4 females	4 males + 3 females
Non-Repetitive Textures	4 males + 5 females	4 males + 5 females

Table 4.7: Number of Interviews included from each conditions

Amongst other things, the interviews were conducted with a view to investigating the interactions between the participant and agents. The analysis particularly focused on the participant's impressions of the agents even though the only behaviours modelled in the agents was the ability to walk from one end of the street to the other while avoiding collisions with the participant and each other. The following research questions were covered:

- How do participants interact with agents without explicit instructions?
- Why do some participants avoid interacting with agents?
- What are participant's expectation of a responsive agent?
- What were the realistic responses experienced by participants to the agents, even though the agents displayed very minimal behaviours?

Generally participants reported that the agents lacked an emotional dimension, realistic gaze behaviour, gestures, variety in behaviour and natural body movement: "*Some people, their body posture might change slightly, if they acknowledge somebody. I cant really explain it. I am not an expert but... I just noticed these things.* (P31121)". Participants also indicated that a lack of agent-agent interaction in the virtual street would have added to the behaviour fidelity of the agents: "*You expect people not to do the same action like walking ahead. I mean they could stop at the window of a shop, chat with each other or smile probably* (P31221)". The findings presented in the following section focuses on participant-agent interactions and participant responses.

4.2.5.1 Participant-Agent interactions and expectations

Participants reported trying to interact with the agents by first approaching the agents, standing in front of them and trying to stop them or bump into them with the expectation that the agent would respond or interact. These were some of the expected forms of interactions anticipated prior to running the experiments, however, twelve participants reported expecting responses to more subtle forms of interactions such as staring or looking at the agent. Most participants tried to interact with the agents in order to see how their actions would cause the agents to react. However, one participant got *lost* in the virtual street and wanted to ask for directions. Whatever the form of interaction, the participant had preconceived expectations of how the agent would react.

Eleven participants expected the agents to respond to their presence using a number of behavioural cues including eye gaze, facial expressions and bodily cues. This was especially the case when the participant made an attempt to interact with the agent.: "*If you see a person, you might like, it will show like on your face, you might like acknowledge that. You might, if you are a guy and you are walking on the street and you look at a good looking girl, you might smile or something, right? Or if you see like a tramp or something, you might have a mournful look on your face, you feel sorry for him or something, so its just expression of feedback* (P31121)". However, ten participants did not even attempt to interact with the agents because they did not expect the agents to respond either because the agents looked computerised or because of prior experiences. Five participants did not interact with the agents because they did not need to or because of the social context provided did not allow them to. Three participants did not interact with the agents for fear of a negative reaction.

4.2.5.2 Non-responsive Agents: The participant perspective

The most significant missing element pointed out by participants were the agent's lack of responsiveness to the participant's behaviour (14) and acknowledgement of the participant (13): "*They were walking in a very mechanical way and they weren't even looking at me and when I bumped into them, they had no feeling, they didn't say anything, they didn't even move their heads* (P22221)". This lead two participants to report that they treated the agents more as virtual objects. The lack of responsiveness led a further two participants to feel that their behaviour in the virtual street had no impact on the agents' behaviour while a lack of acknowledgement, caused through a lack of participant-agent eye contact, made eight participants feel ignored and one participant feel unimportant.

Even though a lack of agent responsiveness resulted in minimal participant-agent interaction, fourteen participants attributed this to the agents being busy, preoccupied and purposeful: "*I just thought that they were just busy. You know that they were going somewhere. And I wasn't part of the process. And I didn't find that threatening or annoying or anything, it seemed quite natural* (P32111)". One participant perceived the agents' non-responsiveness as being due to the agents' arrogance. A lack of responsiveness lead five participants to evaluate the agents as cold and unfriendly which in turn lead these participants to avoid situations which would result in an interaction with the agent: *"Cold and unfriendly and not liking any contact or maybe I wasn't making contact with them, I deliberately avoid to make a contact with them.* (P12111)". As already mentioned, in Section 4.2.2.5, the agents were programmed to avoid collision with the participant, however, one participant viewed this action as a sign of unfriendliness when trying to interact with the agent.

4.2.5.3 Participant responses and perceived agent attributes

Most participants recognised the agents as unexpressive or neutral, however, some participants interpreted the simple behaviours modelled in the agents to signify different states. In addition to being described as cold, unfriendly and indifferent, the agents were also described as annoyed, arrogant, grumpy, indifferent, preoccupied and purposeful.

One of the participants perceived different emotional states in different agents in the virtual street: *"The Saudis were walking, just going about their business quiet calmly. And then the lady I felt was a bit annoyed when I bumped into her* (P21111)". The same participant went on to report how he identified with the agents when the experiment was coming to an end: *"It was funny, in the end when the light went down, and I had this fear of - I had to get out. Trying to find the way and I was lost. And then I saw a woman go by, and I wanted to stop her but I thought she was running as well and I had this impression/feeling that everybody was doing the same as I was* (P21111)". In addition to fourteen participants simply attributing a state of preoccupation to the non-responsive agents, another participant explained their behaviour in terms of the agent's perceived emotional state (grumpy). A further five participants avoided interacting with the agents because in addition to perceiving the agents as cold and unfriendly while one participant did not interact with the agents because they couldn't identify common interests with the agents. Three participants felt that the agent did not want to interact since the agents avoided eye contact and was non-responsiveness.

More interestingly, 33 out of the 34 participants reported feeling various degrees of realistic responses to interactions with the agents in the virtual street independent of experimental condition. These responses ranged from feeling anxious in the presence of the agents to having memories of a similar emotional situation from the physical world during the experiment. For instance, one participant reported an urge to ask agents for directions when he felt lost. Another participant reported that she did not interact with the agents for fear of a negative reaction (Section 4.2.5.1). Three participants felt hesitant to break social norms and bump into the agents even though they wanted to test the agents. Ten participants made a deliberate effort to avoid colliding with agents: *"From my experience before everybody I tried to bump into, everyone I tried to have contact with, they all moved to the side. I tried to do the same thing with her but I kind of stopped and it seemed like she did come at me. And then I jumped back and I was shocked that it had happened. And then again I laughed because it was a bit silly being scared of something that doesn't exist* (P31211)". Five participants even expected the agents to apologise either vocally or using gestures in a situation that can be described as virtual pedestrian rage: *"I remember bumping into someone. And I got a bit annoyed because he didn't turn around and say sorry* (P51121)".

4.2.6 Summary

The experiment on visual realism was designed to investigate if the visual consistency within a scene enhanced participant responses with respect to self-reported presence. In addition the experiment sought to test a new animation library, PIAVCA, and was used to explore the use of physiological measures of participant responses. The main statistical result of the experiment was that the condition with repetitive textures and higher fidelity agent representations produced a lower reported presence response from the other three conditions. This meant that the original hypothesis, that the level of visual realism needs to be consistent is not fully supported. However, it is in keeping with results obtained in the first preliminary experiment. In addition, to there being a need for consistency between visual realism and behavioural realism of a virtual human, there appears to be an interaction between the levels of visual realism in a VE. There are many different types of visual properties that could be changed, however, it is difficult to make the judgement about whether the visual appearance of different elements is consistent.

A final point to make is the highest significant subjective measure in the study was the participant's perceived sense of realism with respect to the behaviour of the agents. The more a participant perceived the agents behaviour to be real, the higher their reported sense of presence. This strengthens the views reported in previous studies (Slater and Steed, 2002). Further analysis of interviews suggested that participants had a number of expectations of how the agents should behave especially with respect to acknowledging the participant. For instance, the most significant missing cue reported by participants was the lack of gaze behaviour in the agents. However, despite the minimal behaviours implemented in the agents and independent of the visual appearance of the agents, participants still reported a wide variety of perceived states in the agents and realistic responses to the agents. This is in keeping with the phenomenon reported by Ushida et al. (1998) where participants reported more perceived states than the actual six modelled in simple spherical agents.

4.3 Discussion

The two preliminary experiments presented in this chapter were inspired by results obtained in early experiments which revealed that virtual humans with minimal behaviours can elicit realistic responses in participants of IVEs. These motivational experiments are not reported within the scope of this thesis, however, they have been reported in Mortensen et al. (2002) and Freeman et al. (2003).

The first preliminary experiment on eye gaze and visual realism confirmed findings by Garau et al. (2001) that avatars with inferred gaze behaviours can have a significantly positive impact on participant's self-reported quality of communication. The most interesting interaction effect observed in the experiment reported that in the case of a lower visually realistic avatar, the inferred gaze model had a *negative* effect on participant response. There was a highly consistent pattern of responses amongst many of the response variables that made up the participant's notion of quality of communication. The overall conclusion was that when the avatar was simple, the better (inferred) gaze model did not improve quality of communication, and indeed there seem to be some cases where it may make things worse. In a sense, the non-realistic random gaze model was more dynamic and visually stimulating than the subtler eye movements of the realistic behaviour model. Perhaps following in the Disney tradition (Thomas and Johnson, 1981), individuals in a shared virtual environment need to be *made* to believe that a visually simplistic or cartoonish avatar is "alive" or expressive. This result was especially astounding given that the random gaze model was an averaged out version of the inferred gaze model. However, for the more realistic avatar, the inferred gaze model improved quality of communication. This confirms findings implied in previous studies (Slater and Steed, 2002; Nowak and Biocca, 2003; Tromp et al., 1998), that participants expect more human-like behaviours from more visually realistic avatars. It is also in keeping with Blascovich et al.'s (2002) argument that the behaviour of an avatar is more critical to evoking an emotional response or a higher self-reported presence score in participants. It cannot be assumed that the results found in the first preliminary experiment will hold when applied to other forms of human behaviour within an immersive VR setting. The behaviour module was lacking in other significant nonverbal behaviours such as posture, body movement, gesture and facial expressions (smiling, lip synching, eyebrow lifts and so on). This was apparent in some of the comments from the participants, however, it does open an exploratory question into other nonverbal behaviours. If the results hold true for other behaviours as well, parametric modelling can potential be an inexpensive way to build expressive virtual humans.

Participant expectations and the impact of implementing only minimal behavioural cues in agents were investigated in the second preliminary experiment on visual realism. The statistical analysis of questionnaire data collected in the experiment revealed that the most significant factor to impact participant's responses was the perceived behavioural fidelity of the agent. The perceived behavioural fidelity of agents was *positively* associated with participant's reported presence in the virtual street. This led to qualitative analysis of the interviews with particular focus on participant-agent interactions. The analysis revealed that, in general, participants expected agents to be responsive and acknowledge the participant through a combination of gaze behaviour, facial expression and body language. Yet despite this lack of behaviours, some participants reported responding to the agents as if they were real. Although some participants reported that the agents looked as if they were randomly walking on a virtual street, even more participants reported that the agents looked purposeful and some even attributed or explained the agent's lack of responsiveness as being due to this. Participants also reported that they found the lack of agent's responsiveness as being cold and unfriendly. This is in keeping with the anecdotal evidence reported by participants in the studies conducted by Slater and Steed (2002) in which agents with no associated behaviour were perceived as robotic and cold.

Observations from the preliminary studies and many others reviewed in Chapter 2 have highlighted three main issues in the modelling of convincing and response-inducing agents:

- Increasing either the visual realism or the behavioural complexity of agents in an ad hoc manner is not sufficient to elicit realistic participant responses. It is important to maintain consistent levels of behavioural fidelity with increasing levels of visual realism. Results from the first preliminary experiment suggest that increasing behavioural realism for agents of low visual quality might even lead to a degradation of participant responses.

- The behavioural expressivity of an agent must be modelled in correlation to the context within which the agent is placed. The right level of subtle cues can perform surprisingly well in achieving the desired effect. The challenging aspect in creating a convincing agent is that of representing plausible behavioural cues to depict a perceived psychological state.

- Responses to different events and stimuli in the physical world is complex and varied. This is partially the problem in trying to build a consistent and realistic behavioural model that can be used to elicit a particular response in participants interacting with the virtual human. Participants will often attribute sentience and a psychological state to perceived behavioural cues in the agent despite knowing that the agent is not real.

4.4 Chapter Summary

	First Preliminary Experiment	**Second Preliminary Experiment**
Relevant section	Section 4.1	Section 4.2
Virtual Human	Avatars	Agents
Factors	Photorealism of Avatars Behaviour fidelity (Gaze) of Avatars	Texture quality of the Buildings Photorealism of the Agents
Participants	12 female pairs and 12 male pairs	20 females and 20 males
Environment	Shared empty Virtual Rooms	Virtual Street
Apparatus	ReaCTor HMD	ReaCTor Physiological devices
Software	DIVE Behaviour Module (C Plugin)	DIVE VRPN (C++ Plugin) PIAVCA (C++ Plugin)
Data Collected	Questionnaires Interviews	Questionnaires Interviews Physiological measures
Data Analysed	Questionnaires	Questionnaires Interviews
Publications	Garau et al. (2003) Vinayagamoorthy et al. (2004b)	Vinayagamoorthy et al. (2004a) Gillies et al. (2005) Edlinger et al. (2006) - Featured VE Brogni et al. (2006)

Table 4.8: Summary of the preliminary experiments

This chapter has focused on two preliminary experiments designed to investigate the impact of varying levels of virtual human behaviour fidelity on participant responses. Section 4.1 presented an experiment which explored two aspects of avatar fidelity: visual appearance and behaviour fidelity (gaze). Section 4.2 focused on a second preliminary experiment in which participants interacted with agents with minimal behavioural abilities. A summary of both experiments is presented in Table 4.8.

As mentioned in Chapter 2, in order to create a complete model of high fidelity nonverbal behaviour in virtual humans, a concrete understanding of human behaviour is needed. On the one hand, it is not always necessary to implement complex models of behaviours in agents in order to elicit appropriate realistic participant responses. Results from the second preliminary experiment suggested that although participants expect agents to be responsive and expressive, agents with limited behaviours can elicit significant responses in participants. The findings from the second preliminary experiment also revealed that a wide variety of realistic responses can be elicited in participants through minimal behavioural cues. However, it is highly probable that there needs to be consistency between visual realism and behaviour fidelity of avatars as revealed in the first preliminary experiment. Furthermore, it is important to uncover the key attributes associated with a specific psychological state in order to portray the state successfully through a virtual human.

Chapters 5 and 6 present experiments designed to investigate if virtual humans can utilise posture and body movement in the portrayal of affect and elicit appropriate responses within IVEs. A combination of methods discussed in Chapter 3 was used in the experiments to collect and analyse various participant responses to the virtual humans.

Chapter 5

Experiment: On Posture

This chapter presents an experiment designed to evaluate participant responses to sequences of affective behaviours exhibited by agents in an IVE. Although there has been substantial research into the role of facial expression in the communication of affect, to date, the existing and limited studies on postures have only been conducted in non-immersive settings. As discussed in Chapter 2, earlier studies by Coulson (2004), Darwin (1872), de Gelder (2006), and Wallbott (1998) suggest that postures play a significant role in the communication of affect. In particular, Montepare et al. (1999) and Walters and Walk (1988) argued that postural attributes play an even more important role in contexts where the facial expressions of the individual is not viewable. Furthermore, Argyle (1998), de Gelder (2006), Montepare et al. (1999) and Planalp et al. (1996) argue that incongruent bodily and facial cues might result in ambiguity while judging the emotional state of other individuals. The role of posture in the communication of affect within a social context has not been investigated with respect to virtual humans in an IVE.

This experiment was used to test a parametric model of affective posture in virtual humans using two particular emotional states: *Anger* and *Sadness*. In Chapter 4, the importance of maintaining and prioritising the appropriate levels of avatar fidelity was explored through gaze behaviour and photorealism. In this chapter, even though the primary interest lies in investigating the importance of postural cues, the importance of modelling congruent behavioural cues is investigated through manipulating two different modalities of expression in the experiment: *facial expression* and *postural cues*. A secondary purpose of this experiment was to explore the range of participant responses to a scenario involving agents portraying minimal behavioural cues meant to portray the two emotional states. The participant responses to the agents were measured at several levels: physiological, behavioural (proximal) and subjective.

Section 5.1 discusses the reasoning behind choosing Anger and Sadness as the two emotional states to investigate while Section 5.2 presents the hypothesis behind the experiment. Section 5.3 - 5.5 discusses the experimental setup while Sections 5.6 and 5.7 presents the participant responses collected and analysis of the participant responses respectively. Finally, Section 5.8 discusses the findings from the experiments while Section 5.9 gives a summary of this chapter.

5.1 Justification of emotion choices

The main emotional state of interest in this experiment was Anger due to its association with threat perception and the varieties of possible responses invoked in individuals witnessing an act of aggression (Section 2.5.2). Anger is seen as a negative emotion therefore there is evidence that others focus a heightened attention on individuals displaying the cues to portray anger (Section 2.5.2.1).

Sadness was chosen as a second emotion to investigate in order to disambiguate between any possible effects caused by simply having affective behavioural cues. Either fear or happiness would have been good choices for the second emotion, however, they both had to be excluded from the study since results reported by Coulson (2004) suggest that both states could not be portrayed unambiguously using postural cues. Angry postures were characterised by a forward shift in body weight while sad postures were characterised by a slumped posture.

In addition, the literature review on kinesics (Sections 2.4.3.1 and 2.4.3.2) revealed that Anger and Sadness were two of the most accurately recognised emotional states through postures and body movement (Coulson, 2004; Montepare et al., 1999). Therefore these two emotional states were used as the basis to investigate the role of postural cues in portraying affect.

5.2 Hypothesis

The main interest of the experiment was to investigate if participants would respond according to expectation in reaction to the underlying emotion (Angry or Sad) portrayed by the postures (with or without accompanying facial expressions). For instance, if a participant encounters an agent exhibiting behavioural cues indicative of anger, does the participant respond accordingly? Which cues play a more important role in instigating the response? The experiment was designed to investigate the importance of posture in a situation where the affective state was directed at a focus other than the participant.

The main hypothesis was that a parametric model of posture would play a significant role in the communication of affect in IVE. The associated premises were that:

- Participants would respond to an affective agent, if the agent displayed appropriate affective postural cues.
- Participants would be able to accurately recognise the underlying emotional state of the agent, if the agent displayed the appropriate affective postural cues.
- More participants would be able to accurately recognise the underlying emotional state of the agent, if the agent displayed affective facial expressions and congruent postural cues.

The Null Hypothesis: There will be no differences found in the participant's responses (physiological, behavioural and subjective) to the three conditions described in the following sections.

5.3 Experimental design: Building the scenario

One challenging aspect in creating a scenario involving agents, for experiments in VEs, is in validating the purpose of the agent without informing the participant about the goals of the experiment. The goals of

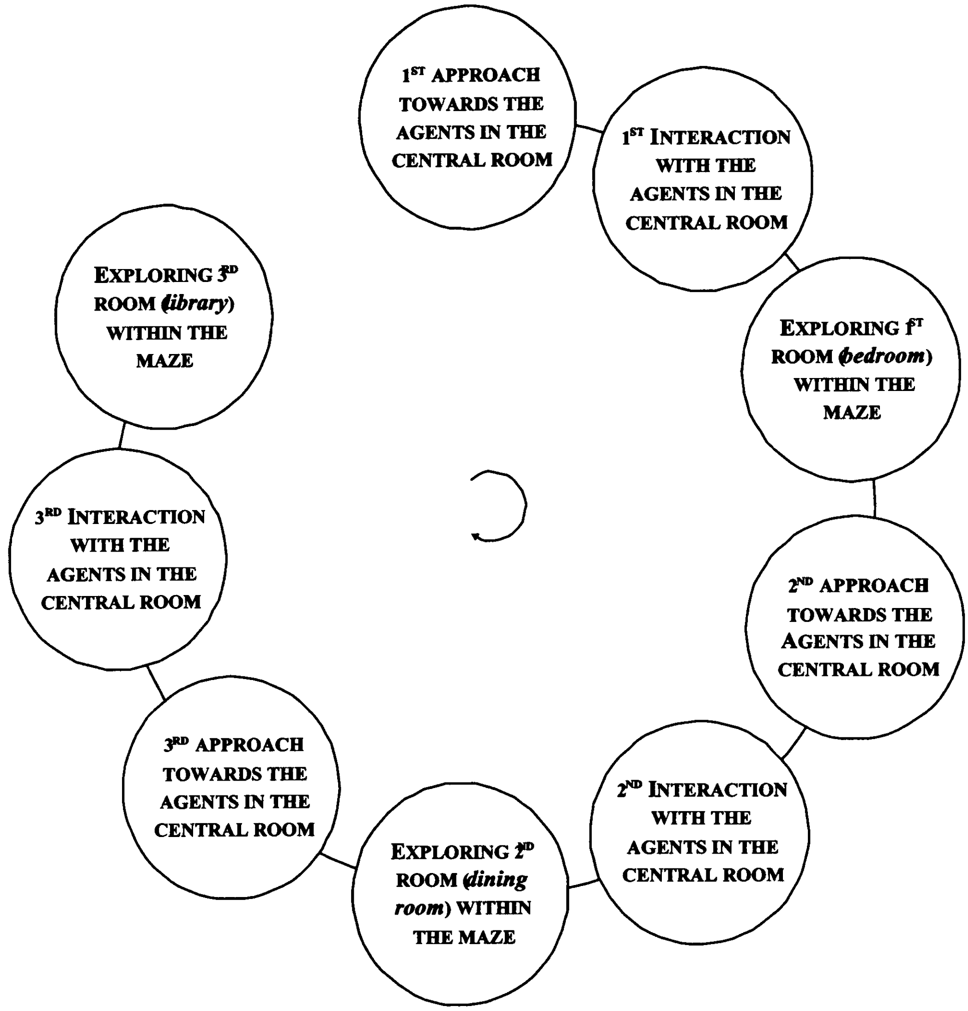


Figure 5.1: Flow of events in the maze

this experiment involved observing and measuring the automatic responses of participants to the agents. In an experiment where participants are asked to evaluate the agents' behaviour, it is reasonable to ask the participant to judge the effectiveness of an agent, however, this changes the behaviour of the participant towards the agent. In this experiment, it was vital to elicit as natural a response from participants given the laboratory setting, therefore, participants were told that the experiment involved interacting with agents in order to obtain directions to explore a virtual maze. Participants were instructed to explore three rooms in the maze, one at a time, using directions given by the agents. Participants were told they would have to explore one room and return to the agents to get directions to the next room.

Throughout this Chapter, one agent is referred to as the '*active*' agent while the other agent is refereed to as the '*passive*' agent. During each of the three periods in which the participant approached the agents, they were seen having a conversation. The verbal conversation between the agents was designed to be muffled (unintelligible) and with neutral vocal tones. The active agent displayed an underlying emotional state (Anger or Sadness) through facial and/or postural cues towards the passive agent during the conversation. The affective state of the active agent was determined in accordance with the experiment condition (Table 5.1). The passive agent portrayed exactly the same neutral state throughout the experiment. A neutral state was defined by ensuring that the agent did not portray behavioural cues designed to be Angry or Sad.

As the participant continued to approach the agents, both agents responded in a neutral manner by turning towards the participant and displaying neutral behavioural cues. The active agent always adopted a neutral state once the participant got close enough to breach the social interpersonal distance (Section 2.4.3.3). The active agent (formerly portraying a condition-dependent affective state) was triggered to greet the participant. When the participant responded and tried to find the right directions to the next room in the virtual maze, the active agent actively interacted with the participant to give directions while the other agent remained a neutral and passive observer to the interaction between the active agent and the participant.

This meant that the participant played two roles in the experiment; the role of an observer to an emotionally charged discussion between two agents followed by a direct face-to-face neutral interaction with the agents. This scenario was developed for two reasons. It allowed participants to observe and form impressions about the perceived affect of the agents without being explicitly instructed. Secondly, an individual's response to an emotional event varies depending on their prior emotional state, personality traits and their perception of factors leading to the stimuli (Section 2.5). This scenario allowed the participant to observe the affective behavioural cues of a virtual character only when it is being directed towards another character in a bystander-type situation.

5.3.1 Design and variables

Two sets (Angry and Sad) of between-groups 2 (*facial expression*) x 2 (*postural cues*) factorial designed experiments were carried out.

	Angry		**Sad**	
	Emotional Face	*Neutral Face*	*Emotional Face*	*Neutral Face*
Emotional Posture	Condition 1	Condition 3	Condition 5	Condition 7
Neutral Posture	Condition 2	Condition 4	Condition 6	Condition 8

Table 5.1: Design of the experiment on participant responses

Due to the manner in which the factors were defined, the control conditions (condition 4 and 8) was repeated twice in the design making one redundant. The repeated condition was the condition in which the active agent displayed a neutral postural and facial cues. The actual set of conditions were therefore:

- *Condition 1:* Angry state portrayed through the use of angry postural cues and angry facial expressions.

- *Condition 2:* Angry state portrayed through the use of neutral postural cues and angry facial expressions.

- *Condition 3:* Angry state portrayed through the use of angry postural cues and neutral facial expressions.

- *Condition 4:* This condition is categorised under the angry state but a brief glance at Table 5.1 reveals that the active agent state is portrayed through neutral postural cues and neutral facial expressions. Conditions 4 and 8 correspond to the neutral (control) condition.
- *Condition 5:* Sad state portrayed through the use of sad postural cues and sad facial expressions.
- *Condition 6:* Sad state portrayed through the use of neutral postural cues and sad facial expressions.
- *Condition 7:* Sad state portrayed through the use of sad postural cues and neutral facial expressions.

5.3.2 Designs, pilots and revisions

The experiment was designed around the objective responses that were being measured: physiological and behavioural (proximal).

Repeated exposure: Early on in the design phase, it was decided that repeated exposure to the agents under the same emotional state would aid in validating participant responses. However, Andreassi (1995) reported that multiple exposures to the same stimuli decreased physiological arousal due to habituation. In addition, too many repeated exposures to the agents would increase the probability of witnessing agents displaying repeated behavioural cues. After repeated testing of the behaviour generation software, a maximum of three exposures to the agents was achieved while maintaining an acceptable level of non-repetitive animation. The three exposures were designed around three participant-agent meeting points (Figures D.2(b) in Appendix D).

Maze designs: Two initial designs for the maze were produced: sequential and centralised (Figure D.2 in Appendix D). There were advantages and disadvantages to both designs. In the sequential maze, there were six different agents (two for each of the three meeting points). The main advantage in the more traditional sequential maze was that the path followed was straightforward. Participants would simply be instructed to get to the end of the maze. A key disadvantage was that six different agents would have to be designed in order to avoid the transportation effect in between meeting points. This was not a problem in the centralised maze since participants would be told to explore rooms which branch off from the main central room (Figure D.2(b)). In the centralised design, the agents would be stationed in the central room at all times therefore the meeting points could be placed at different locations within the central room. An added advantage in the centralised maze was that only two agents were needed; thereby reducing the workload needed in designing or adapting agents. The final VE consisted of a training room leading to a centralised maze.

Technical revisions: The most significant revision made during the pilots was in the behaviour of the active agent. In earlier versions of the agent behaviour, both agents turned towards the participant as soon as the participant got within a certain distance from the agents. This appeared robotic and unrealistic. Some background swaying motions were added to the agents to make their behaviour look less synthetic. In addition it was noticed that the condition-dependent affective facial cues displayed by the active agent were not clearly visible to the participant, since the participant always approached the agents from a side due to the design of the maze. The behaviour of the active agent was modified slightly to include a momentary glance over at the participant in the existing emotional state before both agents turned to face the participant in a neutral state. This had the effect of making the holistic behaviour of the agents more natural. The glance over part of the behaviour made it look as if the active agent became aware of the participant's presence in the room and then signalled this to the passive agent. In addition the active agent was programmed to turn towards the participant slightly earlier than the passive agent in order to avoid the appearance of a robotic turn.

The positions of both agents were modified repeatedly during the pilots in order to allow for maximum participant exposure as soon as the participant entered the room. It was important to ensure that the participant saw both agents at the same time. This was especially important since the analysis of the physiological data would be made under the assumption that the participant saw the agents as soon as they entered the room. Most of the other changes concerned minor aspects of the VE.

- The total polygon count of the maze structure was reduced by a third in an attempt to boost the refresh rate of the VR system.
- It was noticed, during pilots, that participants kept walking through a virtual curtain modelled in front of a transparent sliding glass door in one of the rooms in the maze. A brick wall was modelled on the other side of the glass door to discourage this occurrence.
- The agents had to be remodelled around the joints in order to improve their appearance. This was done by extending the geometry around the joints to allow for bending and twisting movements.

Revisions to the questionnaires: A questionnaire designed to gather the participant's impression of the agent's emotional state were altered to gather more focused participant responses. Participants, in the pilots, found it difficult to understand some of the questions. Initially the questionnaire asked participants to choose one of seven emotional states (happy, sadness, angry, afraid, surprise, disgust and neutral) for each agent. Participants interpreted the question to mean the emotional attitude of the agent towards the participants only. The question was replaced with two questions asking for the participant's impression of each agent's emotional state towards the participant and towards the other agent.

Piloting the task: Finally, the pilots provided an opportunity to try out slightly different versions of the task given to participants. Initially the task was envisioned as one of exploratory similar to the task used in the second preliminary experiment (Section 4.2). Early participants in the pilots were told to ask the agents for directions to three rooms and return after exploring the three rooms. Participants were not given a time limit in which to complete the task since it was important to allow participants to form impressions without feeling the pressure to complete the maze. However, an exploratory task resulted in a wide variation in the duration of the experiment. Some participants found the rooms bare and completed the maze in very little time while others spent too much time in each room exploring. Therefore the task was modified to include a search and count sub-task. Participants were told to explore the rooms and keep a count of the number of trash bins they saw in the rooms. This gave participants a reason to explore the rooms within an acceptable period of time.

5.3.3 Apparatus

Two main sets of apparatus were used. The first was the Trimension ReaCTor system (Section 3.2.3.1). The second set was used to collect data from the physiology monitoring devices using the ProComp+ device (Section 3.2.3.3). Peripheral pieces of equipment included video and audio data recording devices and a file server for backing up the collected data on a daily basis. A Beyerdynamic Diversity Receiver S250 system was used to record audio data through a wireless microphone. The apparatus setup is illustrated in Figure D.1 in Appendix D.

5.3.3.1 Software

The DIVE software with the VRPN and PIAVCA plugins described in Sections 3.2.4.1, 3.2.4.2 and 3.2.4.3 were used to implement the experiment.

Almost all the experiment controller-based events in the maze were handled using the DIVE-Tcl interface through a customised Tk graphical user interface (GUI) (SICS, 2006). As described in Section 3.2.4.1, objects within the VE had a number of DIVE properties which were altered during the course of the experiment to change the state of the object. For instance objects were set to an invisible status, made non-graspable and set to detect collision events with the avatar of a participant. A key feature utilised in the DIVE platform was the handling of the collision events generated and timers through the DIVE-Tcl interface. Scripts were written to bind Tcl procedures to DIVE events (mostly collision-type events), so that bound procedures were called each time these events occurred. In addition agent behaviours and environment animations were controlled by binding Tcl commands to timers, so that bound commands were called each time a timer expired. The DIVE actor attached to a participant was set to register collision events which occur every time the actor gets within a controllable distance of a collision-detectable object. This property was used to trigger events in DIVE which in turn triggered agent behaviours and other environment-based animations such as the activation of automatic doors.

In addition to the core functionality of PIAVCA, a posture loading extension module for PIAVCA was written to animate the postures assigned to the agents. The module was created to enable the experimenter to assign a set of emotion-grouped postures to each agent in the scenario.

5.3.3.2 Posture loading extension to PIAVCA

A posture loading extension to PIAVCA was created to enable the assignment of emotions and appropriate postures to agents. The extension allowed the experimenter to create a number of emotional states and link each state to a number of postures, gestures or other animations with an associated intensity factor. A particular posture could be mapped to more than one emotional state. For the purposes of this experiment, all bodily cues had the same intensity factor and was mapped to one emotional state. In principle, each agent could be assigned a number of emotional states to allow for emotional blends. However, in this experiment, each agent was assigned one emotional state at a time. The active agent was assigned a condition-dependent emotional state during the periods of the experiment in which the active agent was seen interacting with the passive agent. During the interaction with the participant, the active agent was always assigned a neutral emotional state. The passive agent was assigned the neutral state throughout the experiment.

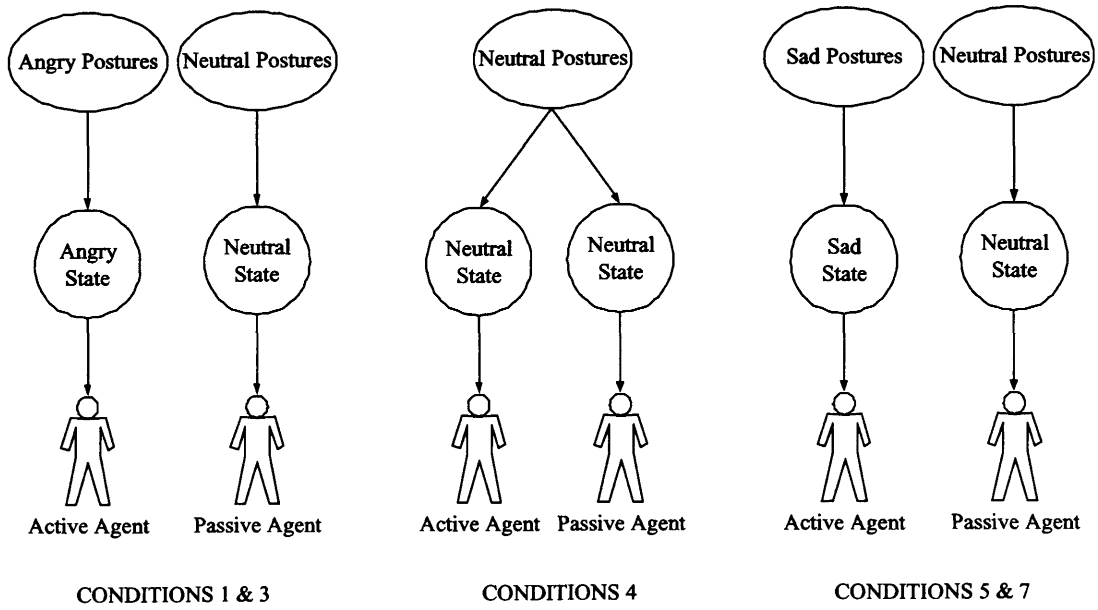


Figure 5.2: Assignment of posture to condition-dependent emotional states in the agents

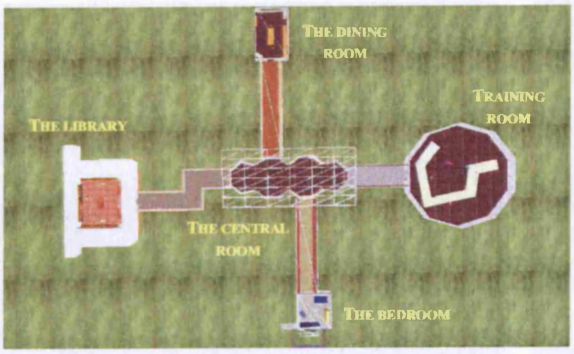
In addition to a set of basic postures to represent the neutral state, two sets of basic postures were adapted from Coulson (2004) for Anger and Sadness. These were mapped to agent names using the posture loading module. The module used a base posture and a set of basic postures to create a much larger interpolated set of postures to represent the corresponding emotional state. The interpolation algorithm used was similar to the one reported by Johnson (2003). Each new posture in the interpolated set was produced by assigning a random weight from a uniform distribution to some or all of the basic postures and summing the resulting poses (joint quaternions). The resulting postures were then put on a PIAVCA motion queue to achieve a smoothly blended animation from one pose to the next.

5.3.4 The Virtual Environment

An elaborate VE was created using 3D Studio Max (Figure D.3). A virtual maze was structured around a central room with eight automatic sliding doors; three of which were activated sequentiality and led via long corridors to one of three rooms - a bedroom, a dining room or a library. The maze was designed with two key features; an open-plan central room allowing participants to choose a comfortable participant-agent distance and long inter-connecting corridors between the central room and adjoining rooms (Figure 5.3(a)). Jang et al. (2002) found that GSR and HRV depicted an arousal of participants exposed to virtual environments and that such measures generally returned to normal over time. The long corridors were designed as a feature in order to allow ample time for the physiology of the participant to return to the tonic level thus minimising a cumulative effect.

5.3.4.1 Training room

In addition, a virtual training room was designed and attached to the entrance of the maze. The training room was used to acclimatise participants to viewing and navigating the VE. A set training program



(a)

(b)

Figure 5.3: a) Top-down view of the virtual environment; b) A participant in the training room.

was administered by an experimenter to ensure all participants were given the same level of information about the capabilities afforded by the VR system. Participants were asked to follow a set of single-digit three-dimensional numbers sequentially from 1 to 5 using a mini-joystick on a wireless wand tracker (Figures 5.3(b) and D.5). Once they reached the number 5, participants were asked to navigate to the start of a yellow winding path in the training room (Figure 5.3(a)). The extensive training in navigation was given because anecdotal statements in the past have suggested that participants often find navigation in long closed spaces difficult (Mortensen et al., 2002). At the end of the training, participants were stationed in front of an entrance door which led into the virtual maze.

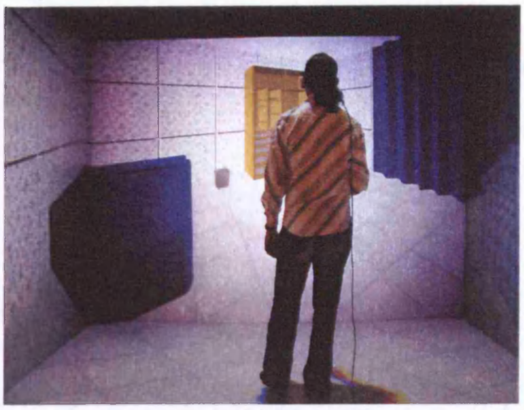
5.3.4.2 The Maze

The training room was connected to a central room with eight possible doors (excluding the entrance door). Three of the eight doors were connected to three adjoining rooms while the other five doors acted as decoys. Each of the eight doors in the central room was labelled with a number (1-8). Doors 4, 1 and 6 were connected through a corridor to a bedroom, a dining room and a library respectively. An environment mimicking a gallery or museum would have sufficed but this may have introduced a confounding factor in which some participant may constantly compare and contrast the virtual environment with a real gallery or museum. A maze allowed for lower levels of visual fidelity.

The central room was the space in which all participant-agent interplays took place. It comprised of four sections welded to form one large open-planned space. It allowed participants to navigate freely in whichever direction best suited them while approaching the agents in one of the meeting points. There were three meeting points (labelled 3, 5 and 7 in Figure 5.5) in the central room; all of which were located in the two midsections of the central room. The meeting points in the central room were located at positions so that the maximum length along the midsections of the central room was utilised. This allowed for longer recordings of the participant's physiological responses while approaching the agents.

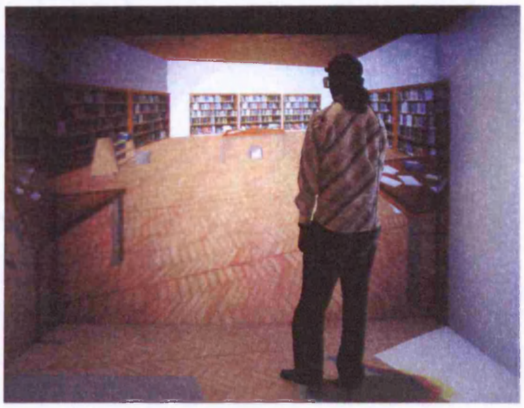
The door leading from the training room into the central room was activated at the start of the experiment. An activated door refers to one which was programmed to open if the participant navigated to a close distance. The appropriate door leading to an adjoining room was activated after each of the three meetings with the agents. The other doors remained disabled in order to force all participants

(a) The central room



(b) The bedroom

(c) The dining room



(d) The library

Figure 5.4: Snapshots of a participant visiting the virtual maze.

to follow the same general pathway in the maze (Figure 5.5). In addition, only the room through the correct door was rendered visible. For instance, in the first meeting with the agents, the participants were advised to go through door 4 into the bedroom. Correspondingly only door 4 leading to the bedroom was activated and only the bedroom was rendered visible. The dining room and the library were set to remain invisible through using property setting functionalities in DIVE. This was to avoid rewarding participants with a room to explore should the participant decide to disregard the directions given by the agents and follow an inactive door. This procedure also had the added advantage of loading only necessary virtual models in the system.

There were five groups of DIVE events used in the experiment which acted either as a trigger for an attached Tcl script or merely as indicators for data logging purposes;

- **Experiment clock indicators:** There were four timer events triggered by the experimenters and logged in the data files. These events were used to calculate the time spent by each participant during different parts of the experiment. The duration of the experiment was calculated using DIVE events triggered at the start and the end of the experiment. The start indicator event also triggered the activation of the entrance door to the central room in the maze.

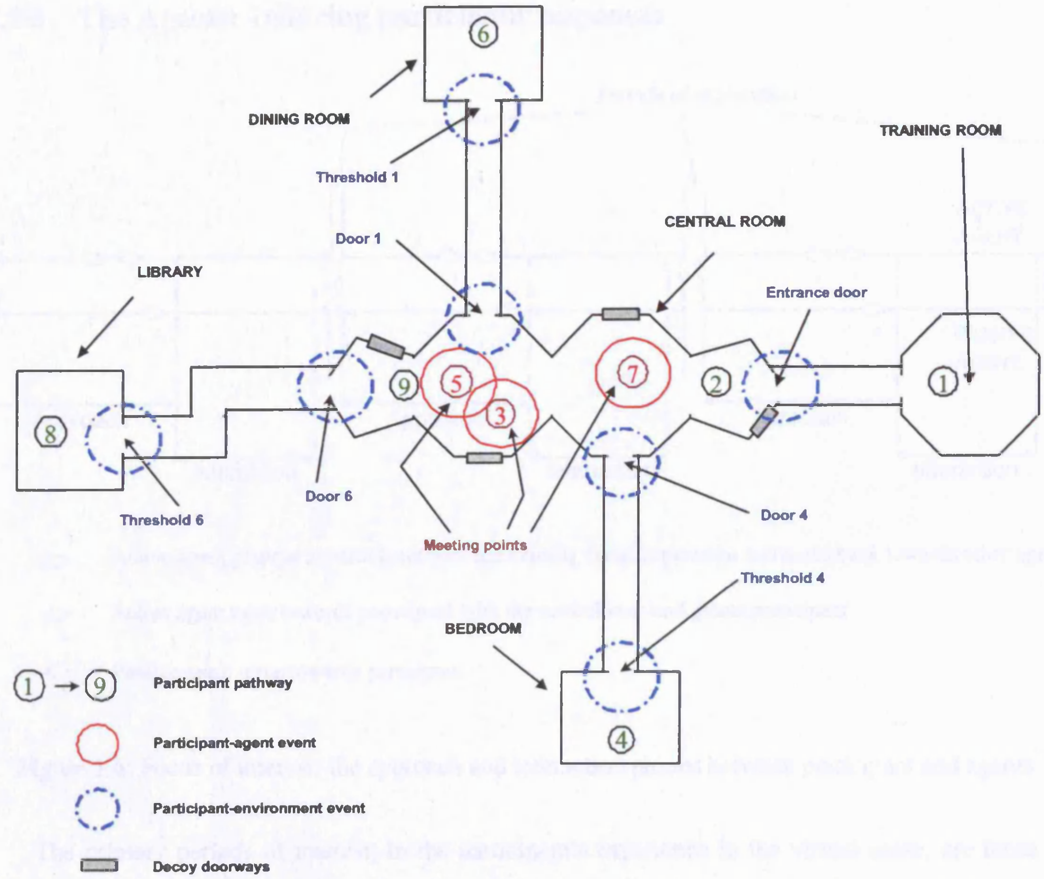


Figure 5.5: Final layout of the maze with illustrations of some events triggered in the environment

- **Door animation triggers:** An activated door was triggered to open if the participant got within 1.5 metres of it. The doors closed after 10000 milliseconds.

- **Room threshold triggers:** The doors in the environment led to a long corridor which in turn led to an adjoining room with a doorway that could detect collisions. These acted as threshold triggers and indicated when the participant enters/exits the room (Figure 5.5). The triggers were attached to a Tcl script which caused the agents to start interacting with each other in readiness for the participant's next arrival in the central room.

- **Agent proximity triggers:** These triggered the second phase of the participant-agent interplay - *interaction.* They occur when the participant gets within the social distance of the agents at 3.25 metres (Hall, 1969). This caused the agent to stop their conversation and behave in a manner that portrayed awareness of the participant. The emotional active agent was set to a neutral state in readiness for the participant-agent interaction. The correct door, leading to the next room, was activated as soon as the participant started the interaction phase.

- **Agent vocal response indicators:** There were a set of 39 possible vocal responses for each meeting with an active agent which gave a total sum of about 117 vocal response events. These events acted as markers in the data log files.

5.3.5 The Agents: Inducing participant responses

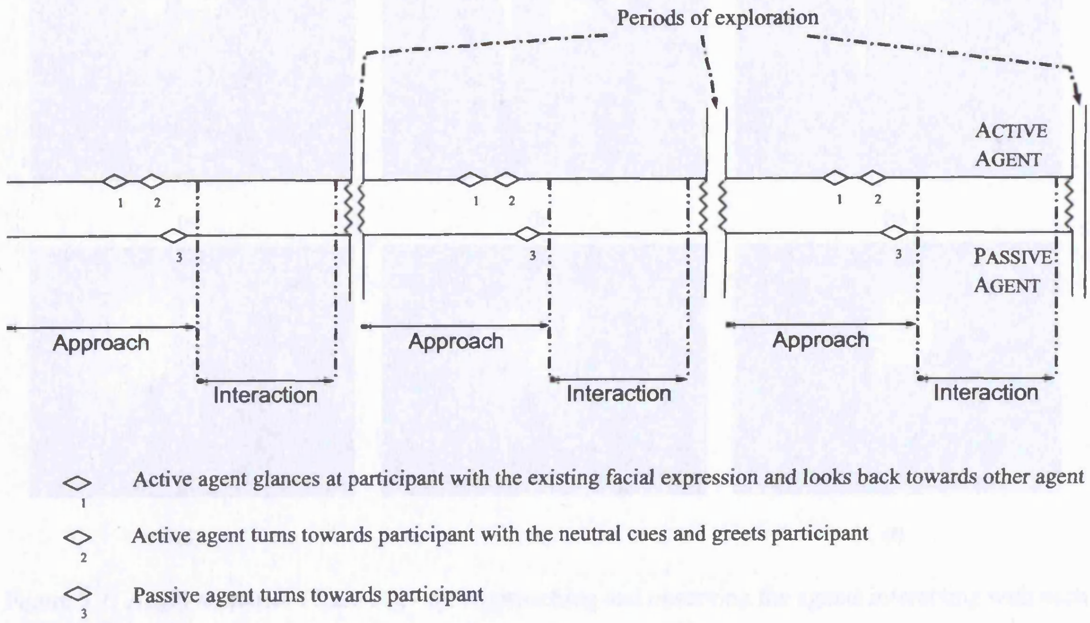


Figure 5.6: Focus of interest: the approach and interaction phases between participant and agents

The primary periods of interest, in the participant's experience in the virtual maze, are those in which the participant was in the central room with the agents. These periods can be divided into two groups: three periods of *approach* and three periods of *interaction*. An approach period is always followed by an interaction period which is then followed by a period of exploration (Figure 5.1). The approach and interaction periods occurred in the central rooms while the period of exploration occurred in one of the adjoining rooms in the maze.

The *approach* periods consisted of the times during which the participants observed the agents having a muffled conversation. The point at which the participant first saw the agents is assumed to the same as when the participant entered the central room. During the approach periods, the participant played the role of an observer. The active agent was seen portraying nonverbal cues meant to portray one of the condition-dependent emotional states (Anger, Sadness or Neutral) while the passive agent always remained in a Neutral state. Once the participant reached a pre-set interpersonal distance with the agents, the active agent glanced at the participant momentarily (1000ms). The active agent then turned back towards the passive agent and adopted a Neutral state. Afterwards, both agents turned towards the participant at slightly different times. This determined the onset of the *interaction* phases of the experiment.

The interaction periods of the experiments consisted of the durations in which the participant interacted directly with the agents in order find out directions from the agents. Appropriate agent responses were triggered by an experiment controller using a set of pre-recorded vocal responses and accompanying behavioural animations. Each lip-synched vocal response was accompanied with neutral nonverbal envelope animations such as head cocks.

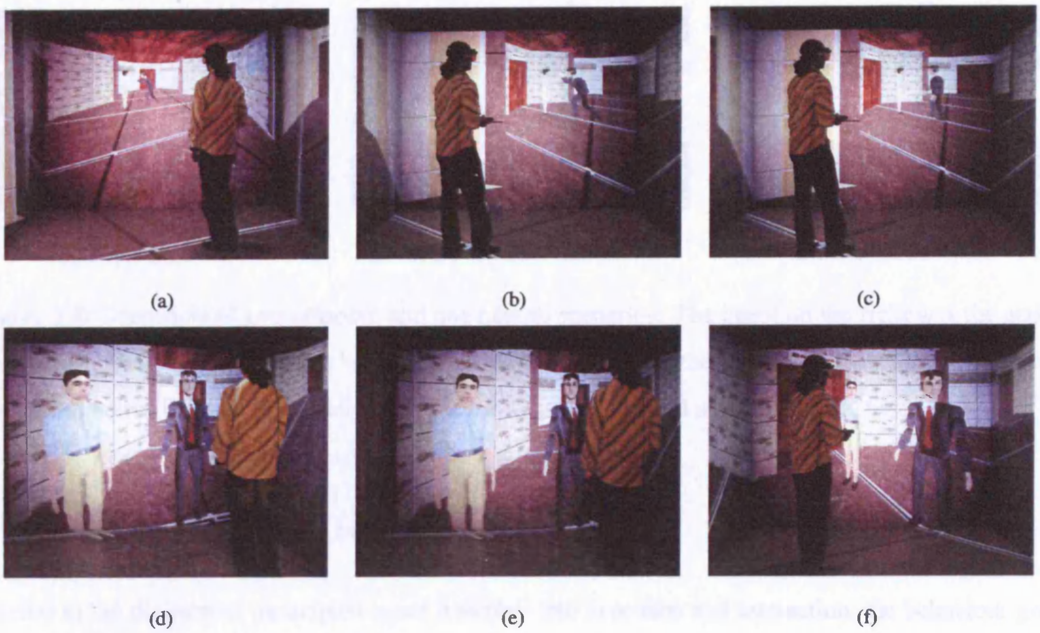


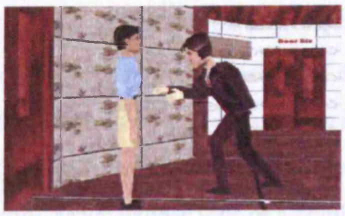
Figure 5.7: Angry scenario: Phase 1 (a - c): Approaching and observing the agents interacting with each other; Phase 2 (d - e): Interacting with the agents to get directions to the next room in the maze

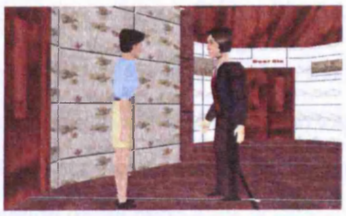
The research goal during both the approach and interaction periods was to determine if participants responded differently to the agents based on the underlying emotional state of the active agent towards the passive agent. For instance, if the active agent is perceived as behaving aggressively towards the passive agent, do participants behave differently towards the agents in comparison to when the active agent is perceived to behave in a neutral manner towards the passive agent?

5.3.5.1 Visual appearance

Two texture-mapped male agents were used to represent the active and passive agents. The agent models were originally adapted from Criterion Software's RenderWare product and used in the preliminary experiments (Chapter 4). The visual appearance of the agents were further enhanced in order to reduce polygon count and smoothen the geometry meshes.

The appearance of the agents were chosen to match the scenario created. In the social context created for the experiment, the active agent was dressed in formal wear (a grey suit) while the other agent was dressed in casual wear (shirt and shorts). The agents were of roughly the same height. Unfortunately the difference in visual appearance with respect to clothing was unavoidable due to the limited number of agents available. However, given the difference in visual appearance, a deliberate decision was made in assigning the role of active agent to the agent in formal wear. In the past, participants have perceived agents to have roles of leadership (higher status) purely on the basis of their visual appearance or enhanced capabilities (Schroeder, 2002). Therefore it was more in keeping with the scenario to assign the role of an active agent to the more formally attired agent.


(a) The angry scenario


(b) The neutral scenario

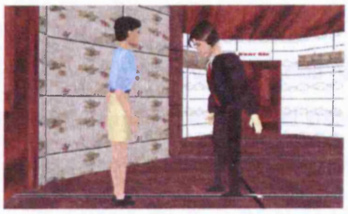

(c) The sad scenario

Figure 5.8: Snapshots of two affective and one neutral scenarios: The agent on the right was the active agent while the agent on the left was the passive agent. While the active agent displays condition-dependent nonverbal cues, the passive agent always displays neutral nonverbal cues.

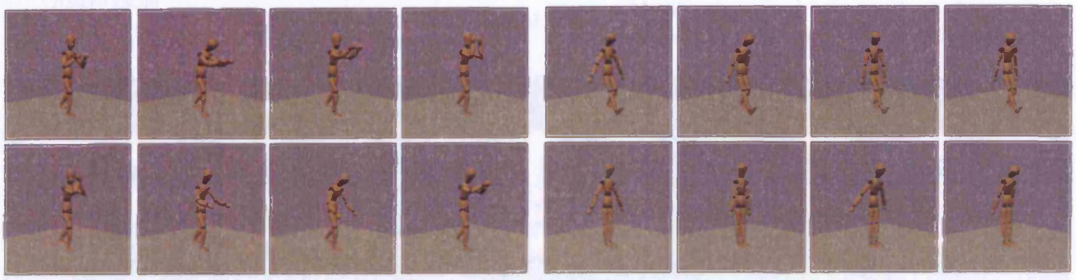
5.3.5.2 Overview of agent behaviour

Similar to the division of participant-agent interplay into approach and interaction, the behaviour generation scheme can be viewed in two parts. Each time the participant entered the central room (three times in one experiment), the agents were seen conversing in the appropriate emotional state until the pre-set participant-agent distance was reached at which time the participant could ask for directions to the next room. The behaviours generated by each agent during the approach periods were aimed towards the other agent and consisted of postural cues and condition-dependent facial expressions. Depending on the condition, the agents changed posture and/or facial expression around a varying time interval. The interval was calculated on the basis of an exponential probability distribution with a mean time of 1000 milliseconds.

Once the participant triggered the interaction phase, a different set of behavioural cues were set in motion to generate neutral behavioural cues aimed at the participant. As soon as the participant sets off an agent-proximity trigger, the active agent stopped focusing on the passive agent and glanced for one second at the participant as if interrupted suddenly. Then the active agent turned back towards the passive agent, adapted a neutral emotional state and turned to face the participant within the next two seconds. The passive agent turned to face the participant as well after a further second. At this point, the experiment controller triggered the active agent to say 'hi' with accompanying non-verbal cues such as a head nod. This behaviour was implemented to gave the impression that the active agent became suddenly 'aware' of the participant, glanced around at the participant (still in the same emotional state) and then turn back to the passive agent, to bring the conversation to a halt, before both agents focused their attention on the participant. During the interaction period of the participant-agent interplay, only the active agent responded to the queries from the participant or took any part in the conversation with the participant. Responses were chosen from a set of pre-recorded messages with accompanying envelope behaviours. The passive agent remained an observer to the conversation and displayed appropriate listening-type animations such as looking at the active agent when it spoke or nodding in agreement. Additionally, slight head movement, regular eye movements and blinking were implemented in both agents.

5.3.5.3 Behavioural generation

During the approach periods of the participant-agent interplay, the active agent was mapped to one of three emotional states depending on the condition while the passive agent was mapped to the neutral state (Figure 5.8). The underlying emotional state of the active agent was portrayed through either postural cues, facial expressions or a combination of the two.



(a) Angry postures (b) Sad postures

Figure 5.9: Affective Postures adapted from Coulson (2004)

The emotional postural cues: One of the research areas that has been little studied is the role of posture in the communication of affect in virtual humans (Section 2.4.3 of Chapter 2). Unlike facial animation, there are no parametric models for the meaningful animation of the bodies of virtual humans. Therefore the main technical challenge encountered in this experiment was gathering data to represent affective postures. Depending on the emotional state of the agent, a set of 32 affective postures, based on Coulson's (2004) research, was associated with each agent using the posture loading extension module. There were then interpolated to get a relatively larger set of affective postures (Section 5.3.3.2). The resulting postures were set onto a motion queue using PIAVCA's core functions at intervals generated using an exponential distribution about a mean of 1000 milliseconds. This created a smooth skeletal animation which was accompanied with facial animations depending on the experiment condition.

Joint	Angry Posture	Sad Posture
Chest Bend†	$20^o, 40^o$	$0^o, 20^o$
Head Bend†	$-20^o, 25^o$	$25^o, 50^o$
Shoulder ab/adduct††	$-60^o, -80^o$	$-60^o, -80^o$
Shoulder swing†	$45^o, 90^o$	0^o
Elbow Bend†	$50^o, 110^o$	0^o
Weight transfer	Forwards	Backwards, Neutral

Table 5.2: Joint rotations based on Coulson (2004); †Positive rotations were forward, ††The neutral position was the arm raised out to the side level with the shoulder. -ve values related to arms above shoulder level '*abduction*' while +ve values related to arms towards the side of the trunk '*adduction*'.

The angry and sad set of postures were adapted from Coulson (2004) while the neutral set of postures were created in 3D Studio Max. The postures adapted from Coulson (2004) were in keeping with observations from earlier studies (Section 2.4.3.1). The neutral set consisted of postures depicting head cocks, arms raised outwards, arms raised slightly upwards and forward/backward leaning poses. None of the postures in the neutral set overlapped with either the angry or sad sets. All the postures designed to represent cues to an underlying affect were only displayed during the approach period of the participant-agent interplay while the agents were perceived to be having a conversation with each other.

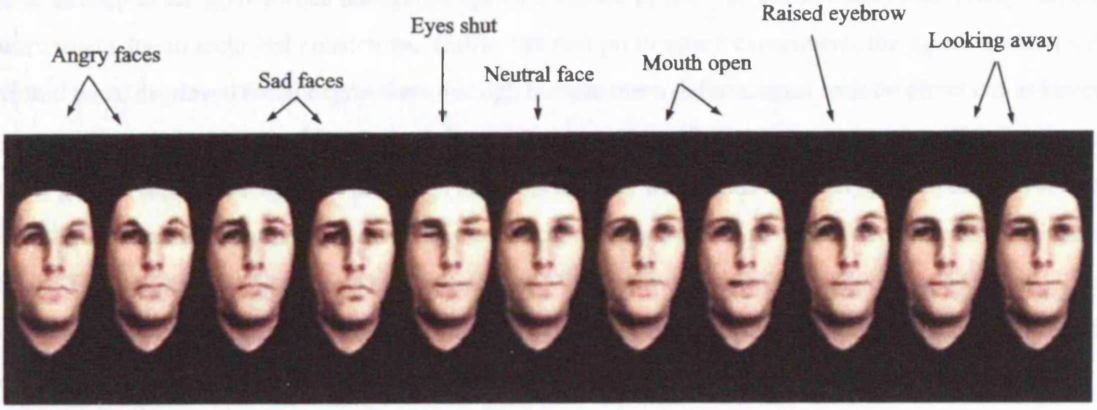


Figure 5.10: Facial expressions created for the active agent

Emotional facial expressions: Facial mesh deformations were created using 3D Studio Max (Autodesk, 2006) and imported into PIAVCA to create facial animations for the agents (Figures 5.10 and D.4). The facial animation functionality, in PIAVCA, was used to create gaze behaviours, blinking and emotional expressions. Two angry and two sad facial expressions were used to create the condition-dependent emotional facial expression for the active agent. No emotional facial expressions were created for the passive agent since it was kept in a neutral and silent state throughout the experiment. Random weights were generated at run time to interpolate between the two appropriate facial expressions and produce a larger set of expressions. Each time the agent changed postures using the posture loading extension module (Section 5.3.3.2), an interpolated facial expression was produced as an accompaniment in conditions 1, 2, 5 and 6. Like the emotional postures, the facial expressions were only displayed during the approach periods of the participant-agent interplay.

Attentive behaviour: Once the interaction periods of the participant-agent interplay was initiated, and both agents had turned to face the participant, Tcl scripts were used to update the agents with the position of the participant's head tracker every second. Even though the active agent was the only vocally responsive agent, both agents responded behaviourally to the participant's head movements by turning to face the participant during the participant-agent interaction. In the second preliminary experiment, one of the flaws pointed out by participant was the lack of eye contact which signals awareness and responsiveness (Section 4.2.5). Furthermore studies conducted by Bailenson et al. (2003) have shown that participants treat agents differently when the agents maintain mutual gaze with the participant.

Gaze behaviour and blink rates: Gaze behaviour was only implemented in the interaction periods of the participant-agent interplay. The first preliminary experiment highlighted the importance of gaze behaviour in avatars especially under a communication context. A cut-down version of the high-fidelity parametric gaze model, described in Section 4.1.5.1, was used to produce gaze behaviours for both agents. In the original gaze model, the avatars were in mutual gaze, only if each participant chose to look 'at' the avatar of their conversation partner and vice versa. The agents used in this experiment were always aware of the position of the participant's head tracker and always faced the participant during the participant-agent conversation. In this case, mutual gaze occurred only when the participant chose to look directly at the agents when the agent's eyes were in the primary position. Most of the other changes were made due to technical constraints. Unlike the first preliminary experiment, the agents used in the virtual maze displayed facial expressions through holistic mesh deformations with no direct control over specific features in the face. This limited the number of distinct 'away' gazes to just two directions; one away gaze to the right and one to the left. Three sets of faces were made for each agent in order to create gaze behaviour and blinking; eyes looking to the right, eyes looking to the left and eyes shut (Figure 5.10). The saccade magnitude was kept constant and there was no control over the saccade velocity. The only factors varied in the gaze model used in this experiment were: inter-saccadic intervals and a limited set of saccade directions.

Like the original gaze model, used in the first preliminary experiment, the gaze behaviour of the agents were dependent on two factors; whether the agent was speaking or listening and whether the agent's eye was in the primary position or not. The agent was set to be in speaking mode if the agent was giving vocal responses and the agent was assumed to be in the listening mode when the participant was talking to the agent. In the active agent, both states were used alternatively while in the passive agent only the listening mode was ever used. In addition to the agents' gaze behaviour, blinking animations were also designed to complement the saccadic behaviour of the pupils. In the active agent, the blinking animation was set as a 3-second looped background facial animation using the eyes shut facial mesh. This resulted in the agents blinking about 20 times in one minute which is about the rate an individual blinks under neutral conditions (Ponder and Kennedy, 1927). The passive agent's blinking animation was set on a 3.5 second blink loop in order to avoid both agents blinking in a synchronised fashion. Even though both blink loops were set on a constant interval, the background animation functionality in PIAVCA produced a loop that was slightly irregular giving the agents a more naturalistic blink rate.

Conversation envelope and lip synching behaviours: During the participant-agent interaction period, in addition to state-dependent gaze behaviour, conversation feedback behaviours were generated as an accompaniment to the vocal responses of the active agent. For instance when the participant asked a yes/no question, the active agent replied vocally and non-verbally by shaking it's head while saying 'no' or nodding it's head to indicate 'yes'. The passive agent would do similar actions but at a subtler speed as if to re-confirm the active agent's response. In addition, a further set of three facial expressions were created for the active agent to create lip-synching animations; mouth open slightly, mouth open fully and raised eye-brows (Figure 5.10).

5.3.5.4 Audio content of the between-agent conversation

During the approach periods, some parts of a pre-recorded conversation were played independent of the nonverbal behaviours of the agents. This gave the impression that the two agents were involved in an active conversation. The conversation itself was masked, emotionless (flat toned) and unintelligible in order to prevent participants from forming an impression about the underlying emotion of the active agent based on the audio content alone.

A number of statements were digitally recorded using the Microsoft Sound Recorder software. The recordings were then manipulated on Sony Sound Forge (Sony, 2006) to remove emotional content in the tone of voice and drastically reduce the intelligibility of the statements by reducing pitch, reversing, filtering and generally transforming the original recordings to mere mumblings. The files were then cut with the original recordings to an extent that allowed participants to recognise it as being a conversation in English. The resulting audio files were converted to uLaw (also known as the Sun Audio) format in order to enable playback in DIVE.

The content of the original phrases were chosen to represent mostly negative content given that the emotional states being investigated were negative, however, it must be stressed that the actual context of the conversation between the agents were unintelligible. Some examples of the phrases recorded were - *'I don't believe it'*, *'you lost it'*, *'directions are important'*, *'this is not good'*, *'very inconvenient'*, *'bother'*, *'I told you to take care of it'*, *'you never ever listen to a word that is said'*, *'it does not work'* and so on. As soon as the participant triggered the interaction period of the participant-agent interplay, the agent-agent audio conversation was halted.

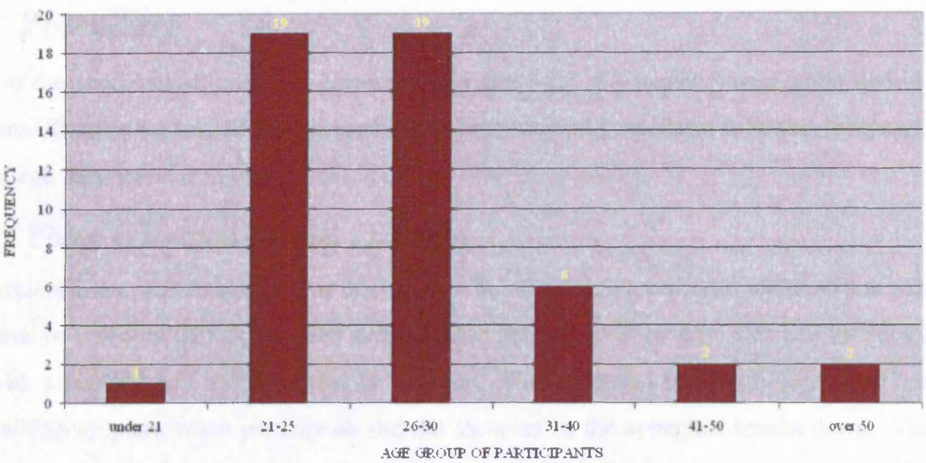
5.3.5.5 Vocal responses to the participant

During the interaction periods of the experiment, a set of pre-recorded audio responses were used to answer participant queries. Each audio response was attached to a Tcl script which triggered appropriate envelope feedback behaviours along with synchronised mouth movements using the facial animation functionality in PIAVCA. When a response was triggered by the experiment controller, the audio response was played back to the participant while the agents display the attached behavioural cues.

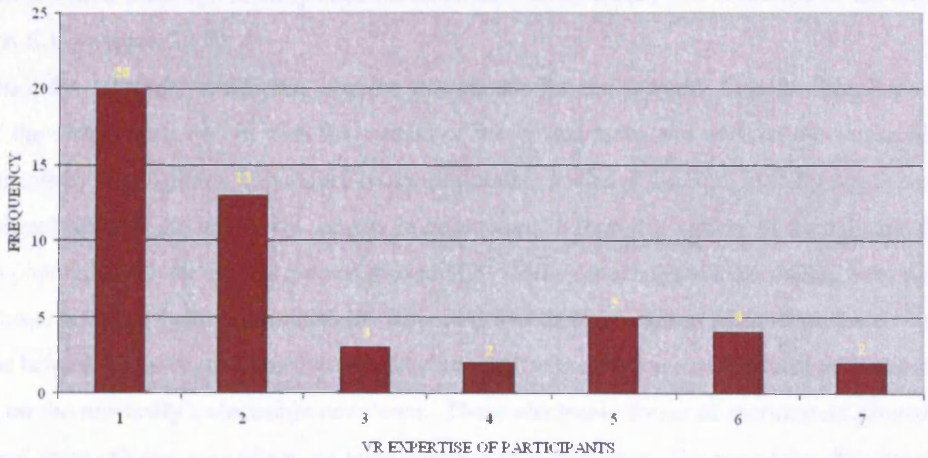
Three sets of responses were used to control the conversations between the participant and the active agent. In the first participant-agent meeting, the participants were directed to the bedroom attached to door 4. During this meeting, the main responses assigned to the active agent were *'hello'*, *'how can we help you?'*, *'only some of the doors can be opened'*, *'door number four should work'*, *'try door number four'*, *'door four'*, *'number four'*. The last four statements were recorded as a preventative measure in case the participant requested to hear the door number again. In the second meeting, the participant was directed to door 1 which led to the dining room. However, the participant had already met the agents once and visited another room. Therefore the active agent in the second meeting were assigned the following statements - *'Did you find it?'*, *'Really?'*, *'What did you see?'*, *'Door number one should open now'*, *'Door one'*, *'Number one'*. In the third meeting, the active agent were given the following statements to help direct the participant to the final room, a library attached to door 6; *'How did it go?'*, *'What was through the door?'*, *'You should try door number six'*, *'Door number six should be active'*, *'Door six'*, *'Number six'*.

A fourth set of general responses was used to give the experiment controller additional responses to choose from should the participant ask an unforeseen question. These responses were also used to deal with events such as the participant speaking in such a low voice that the experiment controller could not hear the participant, the participant asking for directions to the door, the participant asking the active agent task-irrelevant questions, or the participant trying to test the capabilities of the agents by waving the wand in front of the agents etc. The general response set consisted of the following statements - *'yes'*, *'no'*, *'not really'*, *'hmm'*, *'oh?'*, *'I see'*, *'too bad'*, *'oh well'*, *'right'*, *'say that again'*, *'pardon?'*, *'excuse me?'*, *'sorry?'*, *'I did not get that'*, *'louder'*, *'over there'*, *'behind you'*, *'behind us'*, *'I can't answer that question'*, *'we don't know'*, *'I really don't know'*, *'I can't say'*, *'I was not told the task'*, *'want to try that door?'*, and *'Is something wrong?'*. Every vocal response triggered a DIVE event marker which was logged in data files as a marker.

5.4 The participants



(a) Age Group



(b) Experience in VR systems

Figure 5.11: Demographics spread of participants: Age and previous experience in VR systems

In order to eliminate results that were dependent on gender, only male participants were recruited to take part in the experiment. Finally, sixty-seven male participants were recruited for experiments excluding informal preliminary pilots. Eleven participants were treated as paid formal pilots during which many of the changes discussed in Section 5.3.2 were made therefore, the data collected were not used in the analysis. Two participants got ill during the training period of the experiment and had to be exempted from continuing the experiment. Sessions with five participants had to be repeated due to irregularities in the procedure caused by technical issues such as ill-fitted physiological sensors and power supply problems with the ReaCTor. Data collected from forty-nine male participants (seven participants per condition) were used in the analysis. As can be observed in Figures 5.11(a) and 5.11(b)[1], most of the participants were in their twenties and had little VR experience. More than 85% of the participants had very low experience or contact with VR experiences. About 43% of the participants had low game-playing experience while 39% classed themselves had high game-playing experience.

5.5 Procedure

Much of the procedure discussed in Sections 3.2.5 and 3.2.7 of Chapter 3 were replicated in this experiment. There were four distinct phases in the experimental procedures followed: preparation, pre-experiment, during and post-experiment.

5.5.1 Phase I: Recruiting and Planning

Once ethical approval was obtained as discussed in Section 3.2.5, a schedule was drawn to allocate set time slots to potential participants and avoid double bookings. Each time slot was an hour and half long with a planned half an hour break in between. These planned breaks allowed experimenters to run contingency plans when participants did not show up or the apparatus breaks down. The breaks also allowed the experimenters to test the apparatus in between experiment sessions. The 49 sessions (7 batches x 7 conditions) were randomised and mapped to a time slot in the schedule. Each session was coded with a four-digit identity code based on the batch number and condition of the experiment (Section E.1 of Appendix E).

Once the schedule was drawn, a name was chosen for the project: *Lost in virtual space*. The title of the experiments tied in with the context of the virtual maze and masked the true goals of the experiments by playing more importance on the explorative nature of the task. Initially poster campaigns were placed through the university campus in an attempt to recruit a variety of participants from the student population. However, this method proved to be inefficient since the experiments were run during the summer term just before examinations. An email was then sent out to the postgraduate community with the help of the graduate administrator at the university. In addition a recruitment advertisement was placed on the university's electronic newsletter. These electronic forms of recruitment proved to be a faster and more efficient way of getting large numbers of participants. The rest of the recruitment phase followed the same process discussed in Section 3.2.7.

5.5.2 Phase II: Pre-experiment

On the day of the experiment, once the participant arrived, they were welcomed by two experimenters in an anteroom leading to the ReaCTor. As discussed in Section 3.2.7, one experimenter gave the participant

[1]Larger versions of the figures are available in Appendix F.1.

an information sheet (Appendix E.2), a consent form (Appendix E.3) and a number of pre-experiment questionnaires (Section 5.6). While the participant tackled the questionnaires, the second experimenter started the DIVE application and the VRPN server software. Once the VE was loaded, the environment was set to an invisible state in order to darken the ReaCTor. The participant was then invited to step into the darkened ReaCTor after which, the experimenters attached physiological sensors and tracking devices to the participant. The participant was also equipped with a wireless microphone to record their conversation with the agents on video.

The participant was asked to stand still in the ReaCTor in order to record a physiological baseline. The second experimenter started the VRPN client software and checked to see if the physiological data being streamed on a GUI was 'normal'. For instance problems such as inadequate sensor-skin contact were detected in this manner before the start of the experiment. The first experimenter took on the role of an experiment observer while the second experimenter took on the role of an experiment controller (Figure D.1). At this point the experiment observer switched off the lights in the room while the experiment controller triggered a DIVE event to indicate the start of the baseline period in the data log files. The baseline period lasted for a minute and a half, after which the training room described in Section 5.3.4.1 appeared on the walls of the ReaCTor. A DIVE event was automatically sent at the end of the baseline period to signal the start of the training period. The experiment observer administered the navigation training after which the experiment controller gave the task to the participant.

5.5.2.1 The Task

The participant was positioned facing the front wall of the ReaCTor which displayed a door that led from the training room through a corridor to the central room in the virtual maze. Participants were given the same task before commencing the experiment.

> *In a few moments that door will open. It leads to a maze consisting of a central room with doors, which lead to surrounding rooms. You will be able to explore three rooms within the maze, however, which three rooms you explore will be decided by the people in the central room. You can ask them for information to the next room. The people will always stay behind in the central room while you explore the three rooms. After you finish exploring each room, come back to the central room so that you can ask the people which door leads to the next room. Your task is to explore the rooms and remember the number (if any) of trash bins you see. While you are in the maze, we won't be able to talk or respond to you. However, if you want to stop the experiment at any time just say so and we will stop immediately.*

The experiment controller answered any queries the participant had about the task, then instructed the participant to wait until the door opened. The most common question asked at this point was: '*How do I ask for information from the people?*'. The experiment controller always answered: '*In English*'. At this point the experiment controller left the participant in the ReaCTor, checked the physiology signals on the PC monitor, and started the video recorder enroute to the terminal through which the DIVE application can be controlled. The experiment observer took a seat behind the participant and started to note down events that occurred during the experiment.

5.5.3 Phase III: During the experiment

The experiment controller triggered a DIVE event to signal the start of the experiment and open the entrance door leading to the maze. As already mentioned, there were two phases in the participant-agent interplay: the *approach* periods and the *interaction* periods. The approach periods was completely automated while the interaction periods were modulated by the experiment controller.

Each time the participant triggered the start of an interaction period, the experiment controller observed and listened to the participant's side of the conversation in order to trigger the most appropriate response on behalf of the active agent. The experiment observer noted down any gestures, unusual comments made by the participant or technical difficulties faced by the participant during the experiment. Once the participant explored all three rooms in the maze and returned to the central room, the experiment controller triggered a final DIVE event to indicate the end of the experiment. The final DIVE event also darkened the VE. The experiment controller switched the lights on, stopped the video recorder and ensured that the data collected from the physiological sensors was saved. The experiment observer helped the participant to remove all sensors before returning to the anteroom.

5.5.4 Phase IV: Post-experiment

In the anteroom, the participants was given more questionnaires including a presence response questionnaire regarding their experience in the VE (Appendices E.5 to E.8). While the participant filled the post-experiment questionnaires, the experimenters prepared the apparatus for the next participant. When the participant finished the questionnaires, the experimenters carried out a semi-structured interview with the participant (Appendix E.10). At the end of the interview, the participant was debriefed and given a complete disclosure regarding the purpose of the experiment. They were then asked not to discuss their experience with anyone who might be a potential participant in future experiments.

5.6 Measured responses

A number of different subjective and objective participant responses were collected in the experiment. Before the experiment began, in addition to a demographic questionnaire (Appendix E.4), the participants were asked to complete two self-reported trait-anxiety measures, one emotion contagion scale, and an affective perception test. During the experiment, the participant's physiological responses and tracking data were recorded. After the experiment, the participants were administered a presence response questionnaire, an agent's perceived emotion questionnaire, and two state anxiety measures.

5.6.1 Subjective measures

A significant number of questionnaires were administered to the participants before and after the experiment. The questionnaires given before the experiment (phase II - Section 5.5.2) were mostly standard psychological questionnaires designed to measure the participant's characteristics. The questionnaires given after the experiment (phase IV - Section 5.5.4) were designed to measure participant responses to the agents and the VE.

5.6.1.1 Anxiety assessment questionnaires

Two questionnaires were used to assess participant anxiety in daily life (trait anxiety) and after the experiment (state anxiety).

Social Avoidance and Distress (SAD) response scale: The SAD is a 28-item, dichotomous, self report measure designed to assess the subjective experience of distress and avoidance behaviours during social encounters (Watson and Friend, 1969, 1987). High scores on the SAD response scale indicate that the individual experiences negative emotions during social interactions and avoids social situations. Participants were administered the SAD response scale during phase II of the experiment. A modified version of the SAD was administered after the experiment (phase IV) to determine the extent to which the participant avoided the agents. The modified SAD consisted of 26 yes/no questions dealing with the participants social response towards the agents (Section E.8 of Appendix E).

State-Trait Anxiety Inventory (STAI): The STAI is a 40-item questionnaire which provides 20 questions to measure an individual's characteristic trait anxiety (T-Anxiety) and 20 questions to measure state anxiety due to an experience (S-Anxiety) (Spielberger et al., 1983). T-Anxiety was measured before the experiment in phase II using the STAI-T whereas S-Anxiety was measured after the experiment during phase IV using the STAI-S. Each item on the two sub-types of STAI was rated on a 1 to 4 scale. Therefore, scores ranged from 20 to 80 with higher scores indicating higher levels of anxiety. In the physical world, individuals who score high on trait anxiety are more likely to have increased state anxiety during stressful activities.

5.6.1.2 The Affect Perception Test and Emotion Contagion scale

Two additional explanatory variables were collected through questionnaires during phase II of the procedures:

- The participant's self-rated measure of their ability to perceive emotions in others through using the Affect Perception Test (APT) (Gilbert and Coulson, 2002) and
- The participant's measure of their sensitivity to catching other individual's emotions through using the Emotion Contagion (EC) scale (Doherty, 1997).

The APT is a uni-dimensional 14-item questionnaire. Each item of the questionnaire was rated on a scale of 1 to 7 where 1 corresponds to 'strongly disagree' and 7 corresponds to 'strongly agree'. The EC scale is a 15-item questionnaire in which each item is rated on a scale of 1 to 5 (1 = never, 2 = rarely, 3 = usually, 4 = often and 5 = always).

5.6.1.3 Agent emotion assessment questionnaire

In order to determine if participants accurately recognised the underlying emotional states portrayed by the active agent, a simple questionnaire was designed and administered immediately after the experiment (phase IV) before any of the other questionnaires. The questionnaire had three sections dealing with the participant's general impression of the agents, the participant's perception of the emotions portrayed by the active agent towards the passive agents (and vice versa), and the participant's impression of the emotion portrayed by both agents towards the participant. The first item included a set of questions used to check if the agents were animated properly. In each of the two latter questions, participants were asked to choose out of seven emotions: *surprised, afraid, angry, happy, disgusted, sad* and *neutral*. The questionnaire is available in Section E.5 of Appendix E.

5.6.1.4 Presence responses questionnaires

The main subjective response variable of interest was the participant's reported presence as discussed in Section 2.6. This was measured, in phase IV of the experiment, using a questionnaire which included the SUS presence questionnaire (Slater and Steed, 2000) and some questions based on the concept of operational presence (Sanchez-Vives and Slater, 2005). The questionnaire was divided into two parts: the participant's response to the situation as a whole and the participant's response to the agents. The questionnaire is attached in Sections E.6 and E.7 of Appendix E.

5.6.1.5 Semi-structured interviews

Semi-structured interviews were conducted with each participant as the final step in the experiment. In keeping with the style of semi-structured interviews, questions were not posed in a set order rather they were put to the participant as and when the subject came up during the interview. In order to gather as much detail as possible without influencing the participant's responses, binary and leading questions were stringently avoided as was the use of the words *presence* or *breaks-in-presence*. A detailed outline of the content of the interview is given in appendix E.10.

There were four broad areas covered in the interviews: the participant's general impression of what was going on in the maze, the participant's feeling of presence and responses, the extent to which participant had an uninterrupted experience, and the participant's impression of the agents. During pilots it was noticed that participants often did not remember a description of the agents and some participants confused the active agent standing to their left with the passive agent standing on their right. A snapshot of the agents (Figure E.2) were shown to the participants after they had completed all the questionnaires to reaffirm that the correct agent was being discussed.

Initially the participant was asked to describe their experience in the maze from start to finish. This allowed the participant to go into as much details as they liked in describing their experience without too much interference from the experimenter. It also allowed the experimenter to note down areas in the interview agenda that could benefit from more detail or specific parts of the particular participant's experience that could be explored further to uncover interesting findings. If there were any irregularities in the experiment, that the participant mentioned, the experimenter kept a note of it.

The second section of the interview dealt with the participant's sense of presence in the maze through talking about presence with the SUS (Slater and Steed, 2000) and a questionnaire to assess if participants responded or behaved as if the virtual environment was real (Sanchez-Vives and Slater, 2005). Following the complimentary methods of drawing qualitative graphs as introduced by Garau et al. (2004), participant's were asked to draw line, points or levels highlighting their sense of being in the maze and their post-hoc rationalisation of their responses/behaviour in the maze (Figure E.3). This helped participants focus on specific sections of the experience during the discussions. The third section was a natural progression from the second section. Participants were asked for specific events that caused sudden interruptions in their experience. They were also asked to judge the amount of time it took for recovery from these interruptions.

The final section of the interview dealt with the participant's impression of the agents. At this point the experimenter referred to the participant's responses from their completed agent emotion assessment questionnaire (Section E.5) and asked specific questions to uncover the reasons behind their answers. For instance if the participant had indicated that the active agent and passive agent had a working relationship, the experimenter would probe for specific details which led the participant to perceive this. The particular areas of interest covered by the experimenter is summarised in Appendix E.9.

5.6.2 Objective measures

5.6.2.1 Physiological response

Participant's galvanic skin response (GSR), electrocardiogram (ECG) and respiration were monitored and recorded using the Procomp+ device described in Section 3.2.3.3.

5.6.2.2 Proximal response

Bailenson et al. (2001) argued for proximal responses to be utilised as a behavioural measure of the realism of agents. Their study on the maintenance of a version of Argyle and Dean's (1965) equilibrium theory in VEs is in keeping with Reeves and Nass's (1996) premise that individuals assign sentience to agents. The advantage of using proximal responses to an agent in IVE is that tracking data of participants in the system is cheaply available. Time stamped positional and orientation tracking data was recorded through the use of the DIVE VRPN plugin (Section 3.2.4.2). The data was logged along with all the DIVE events triggered in the maze. In addition the positions and orientation of the agents in the maze were also recorded.

5.7 Analysis of responses

5.7.1 Findings from questionnaire responses

Participant responses collected through questionnaires were analysed using the logistic regression method discussed in Sections 3.4.1.1 and 3.4.1.2. Explanatory variables fitted to the regression model included: the participant's age (Age), level of computer literacy (Literate), level of experience with computer programming (Program), level of experience with VR systems (VR), level of experience with video game playing (Game), amount of time spent playing video games per week (Gametime), measure of social avoidance and distress (SAD - measured through the *SAD*), measure of trait anxiety (Trait - measured through the *STAI-T*), measure of affect perception (APT - measured through the *APT*), and measure of sensitivity to catching emotions (EC - emotion contagion as measured through the *EC*). The analysis of the questionnaire responses focused on uncovering factor-related differences in presence and social anxiety responses reported by the participants.

There were significant factor-related results in the responses collected through the SUS ("being there") and "operational co-presence" indicators in the Angry conditions. In the Sad conditions, the "operational presence", "operational co-presence" and social anxiety ("modified sad") indicators yielded significant results. Each of these terms were included in the overall model only if they were significant at the 5% level of significance. In other words, the variable was only included in the model if deleting it significantly reduced the overall fit of the model.

5.7.1.1 Presence response: SUS

In the Angry conditions, an interaction effect was observed between the two factors: type of facial cue and type of postural cue. Separately, the neutral postures and the neutral facial expressions were associated with less reported presence as measured using the "being there" indicator. However, the neutral posture with the neutral facial expression was associated with higher reported presence (Table 5.3). The P values reported in this thesis correspond to deleting the variable from the fitted model. The explanatory variables included in the final regression model included the paticipant's age (Age), the amount of time participants spent playing video games in a week (Gametime), how computer literate the participant was (Literate), their knowledge of computer programming (Program), the level of experience participants had in VR systems and the participant's measure of social avoidance and distress (SAD).

	Deviance χ^2	d.f.	Association	$\sim$ P value
Neutral Posture • Neutral Face	6.97	1	+	0.0082
Age	25.01	1	+	5.70e-007
Gametime	12.95	1	+	3.19e-004
Literate	7.23	1	+	0.0072
Program	12.22	1	+	4.73e-004
VR	14.25	1	−	1.60e-004
SAD	7.43	1	−	0.0064

Table 5.3: Fitted logistic regression for presence ("being there") response in the Angry conditions.

The overall model fitted had a deviance (χ^2) of 51.3 on 18 d.f. At the 5% level of significance, for a model with good fit, the overall deviance (χ^2) would be less than 28.9 at 18 d.f. This indicates that although the individual terms are significant, the overall explanatory power of the model is low, which means that there are important explanatory variables missing. Similar analysis on the "being there" presence response variable for the Sad condition revealed no significant factor-related results.

5.7.1.2 Presence response: The operational definition

Analysis of the operational presence response variable within the Angry conditions did not yield any significant factor-related results. However, in the Sad condition, the neutral postures was associated with higher reported presence (Table 5.4). The explanatory variables fitted to the regression model included Literate, the participant's level of sensitivity to emotion contagion (EC) and the number of times participants played video games in the past year (Game).

The overall model fitted for the operational presence response variable therefore included the type of Posture, the computer literacy of the participant, the participant's sensitivity to emotion contagion and the Game variable ($\chi^2 = 49.4$ on 23 d.f.). At the 5% level of significance, χ^2 has to be less than 35.2 at 23 d.f. for a model with good fit, therefore the overall explanatory power of this model was low as well.

	Deviance χ^2	d.f.	Association	$\sim$ P value
Neutral Posture	7.36	1	+	0.0067
Literate	9.41	1	+	0.0022
EC	6.63	1	−	0.01
Game	11.23	1	−	8.05e-004

Table 5.4: Fitted logistic regression for presence ("operational presence") response in the Sad conditions

5.7.1.3 Co-presence response: The operational definition

The operational co-presence indicator was designed to capture the responses of the participants to the agents. Tables 5.5 and 5.6 depict the significant variables fitted into the overall model, deviances (χ^2), $\sim$ P value and direction of association. Again, the P values correspond to deleting the variable from the fitted model.

	Deviance χ^2	d.f.	Association	$\sim$ P value
Neutral Posture	4.67	1	+	0.0304
Trait	8.86	1	−	0.0029

Table 5.5: Fitted logistic regression for co-presence response in the Angry conditions

	Deviance χ^2	d.f.	Association	$\sim$ P value
Neutral Posture	8.79	1	+	0.0030
Neutral Face	5.73	1	+	0.0167
Game	5.05	1	−	0.0246
Gametime	5.91	1	+	0.0150
VR	7.33	1	+	0.0068

Table 5.6: Fitted logistic regression for co-presence response in the Sad conditions

In both the Angry and the Sad conditions, the neutral postures were associated with an increase in reported co-presence. In the Angry condition, there was a negative association between the trait anxiety of the participant and the reported co-presence. The overall model for the Angry conditions included Posture and the participant's level of trait anxiety ($\chi^2 = 66.3$ on 25 d.f.). At the 5% level of significance, χ^2 has to be less than 37.7 at 25 d.f. for a model with good fit, so the explanatory power of the model is low.

In the Sad condition, as well as the postural cues, the neutral facial cues were also associated with higher reported co-presence. The VR variable was also positively associated with reported co-presence. The overall model for the Sad conditions included type of Posture, type of Facial cue, the participant's level of experience with video game playing (Game), the amount of time the participant spent playing video games per week (Gametime), and the participant's level of experience with VR systems ($\chi^2 = 26.09$ on 22 d.f.). Unlike the Angry conditions, the overall model for the Sad condition is of a good fit.

5.7.1.4 Social avoidance and distress

	Deviance χ^2	d.f.	Association	$\sim$ P value
Neutral Posture	9.69	1	+	0.0018
Neutral Face	25.29	1	−	4.93e-007
SAD	9.86	1	+	0.0017
Trait	4.38	1	+	0.0363
APT	13.08	1	+	2.98e-004
Age	29.08	1	+	6.95e-008
Program	9.78	1	+	0.0018
Gametime	5.35	1	+	0.0207

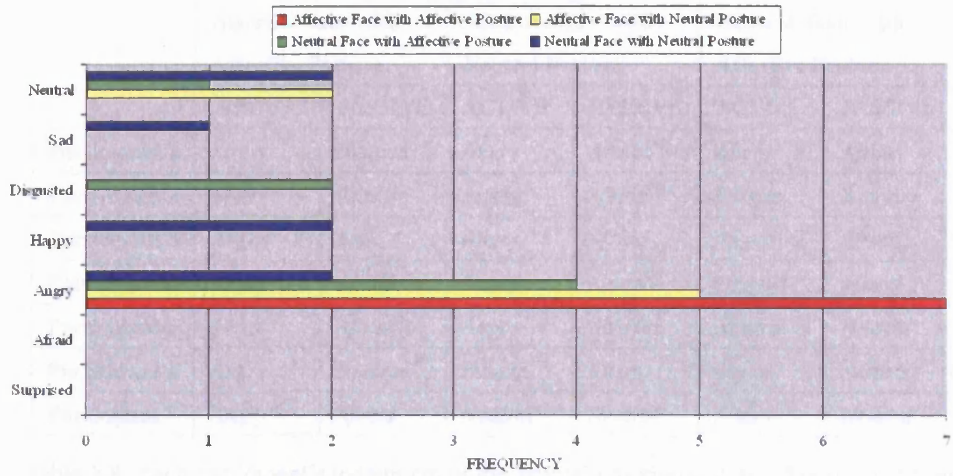
Table 5.7: Fitted logistic regression for modified SAD response in the Sad conditions

The final indicator "modified sad" measure evaluates the participants' level of self-reported social anxiety in response to the virtual characters. In the Angry conditions there are no significant factor-related results. In the Sad condition, the neutral postures are associated with increased social anxiety while neutral facial cues are associated with reduced social anxiety. Other variables in the Sad condition that were associated with an increase in social anxiety included the age of the participant, their knowledge of computer programming, the participant's level of social anxiety experienced by the participant in the physical world, the participant's APT score, their trait anxiety level and Gametime. The overall model fitted had a deviance (χ^2) of 40.6 on 19 d.f. At the 5% level of significance, χ^2 has to be less than 30.1 at 19 d.f. for a model with good fit, so the overall explanatory power of the model is low.

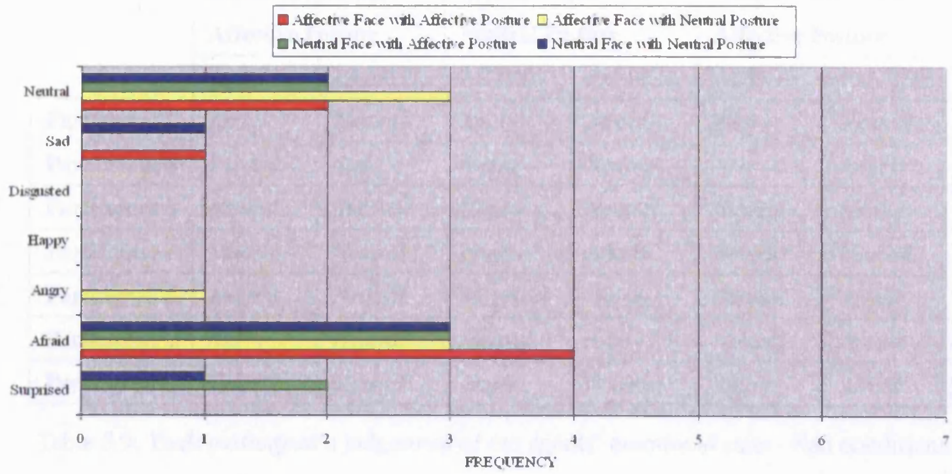
5.7.1.5 Accuracy in recognising underlying emotion

There were two primary aspects of the agent's behaviour that participants were asked to judge using the questionnaire discussed in Section 5.6.1.3 (Appendix E.5). The first dealt with the participant's judgement of the agent's underlying emotional state towards the other agent. The second dealt with their judgement of the agent's underlying emotional state towards the participant.

As can be observed in Figure 5.13(a), in the Angry conditions, all seven participants who were in the conditions with affective facial expressions and affective postures accurately judged that the active agent's emotional state towards the passive agent was angry. Four of these participants judged the passive agent as being afraid while one participant judged the passive agent as being sad. This is in keeping with their judgement of the active agent as being angry. The remaining two participants judged the passive agent as being neutral which is not inconsistent with their judgement of the active agent. Five participants in the condition with affective facial expressions only, judged the active agent's emotional state towards the passive agent as angry while the remaining two participants judged the emotional state of the active agent as neutral. Three participants in this condition judged the passive agent's emotional state towards the active agent as afraid, while three participants judged the passive agent as being neutral towards the active agent. One participant who thought the active agent was neutral thought that the passive agent was angry in this condition. Four participants in the condition with affective postural cues only, accurately



(a) Active agent



(b) Passive agent

Figure 5.12: Participant's judgement of the agents' emotional state towards each other - Angry conditions

judged the active agent as being angry towards the passive agent while two participants thought the active agent was disgusted at the passive agent and one participant thought the active agent was neutral. The passive agent was judged as being afraid by three participant, surprised by two participants and neutral by the remaining two participants. Again the emotional state of the passive agent is judged in keeping with the emotional state of the active agent (Table 5.8).

In the Sad conditions (Table 5.9), not even a single participant in the conditions with affective behavioural cues (facial or postural) identified the active agent's emotional state towards the passive as sad. Even in the conditions where the active agent displayed congruent and affective behavioural cues, four participants judged the active agent as neutral, two participants judged the active agent as angry and one participant judged the active agent as disgusted at the passive agent. Figures representing participants' judgement of the agents' emotional states in the Sad conditions is attached in Figure F.13 (Appendix F.6). However, in keeping with the results observed in the Angry conditions, the participants

	Affective face with Affective Posture		**Affective face with Neutral Posture**		**Neutral face with Affective Posture**	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE	ACTIVE	PASSIVE
Participant 1	*Angry*	*Neutral*	*Angry*	*Afraid*	*Angry*	*Afraid*
Participant 2	*Angry*	*Afraid*	*Angry*	*Afraid*	*Disgust*	*Surprise*
Participant 3	*Angry*	*Sad*	*Angry*	*Afraid*	*Angry*	*Afraid*
Participant 4	*Angry*	*Afraid*	*Angry*	*Neutral*	*Disgust*	*Afraid*
Participant 5	*Angry*	*Afraid*	*Angry*	*Neutral*	*Neutral*	*Neutral*
Participant 6	*Angry*	*Neutral*	*Neutral*	*Angry*	*Angry*	*Surprise*
Participant 7	*Angry*	*Afraid*	*Neutral*	*Neutral*	*Angry*	*Neutral*

Table 5.8: Each participant's judgement of the agents' emotional state - Angry conditions

	Affective face with Affective Posture		**Affective face with Neutral Posture**		**Neutral face with Affective Posture**	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE	ACTIVE	PASSIVE
Participant 1	*Angry*	*Neutral*	*Angry*	*Afraid*	*Angry*	*Neutral*
Participant 2	*Disgust*	*Sad*	*Angry*	*Neutral*	*Neutral*	*Angry*
Participant 3	*Neutral*	*Surprise*	*Happy*	*Neutral*	*Neutral*	*Neutral*
Participant 4	*Neutral*	*Neutral*	*Angry*	*Afraid*	*Neutral*	*Neutral*
Participant 5	*Neutral*	*Neutral*	*Surprise*	*Happy*	*Disgust*	*Afraid*
Participant 6	*Angry*	*Neutral*	*Neutral*	*Angry*	*Neutral*	*Afraid*
Participant 7	*Neutral*	*Neutral*	*Angry*	*Neutral*	*Angry*	*Afraid*

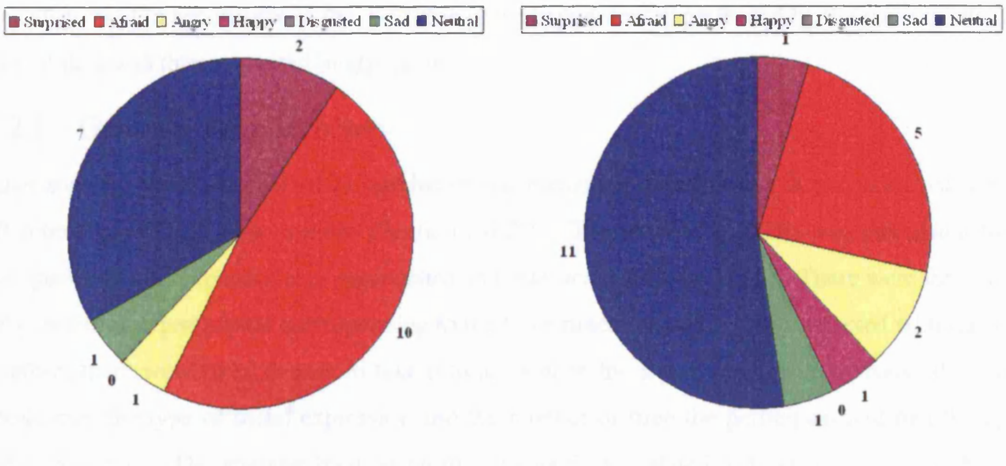
Table 5.9: Each participant's judgement of the agents' emotional state - Sad conditions

judged the emotional state of the passive agent such that it was complementary to the perceived emotional state of the active agent.

In fact, one of the most interesting outcomes from the analysis of the participant's judgement of agents' emotional state was concerning the emotional state of the passive agent. The passive agent was designed to always portray the same behavioural cues throughout the whole experiment, however, only 20 out of the 49 participants perceived the passive agent to be neutral (Figure 5.13).

A similar effect was noticed in the Neutral conditions which is surprising since other there were no affective behavioural cues displayed by the agents throughout the whole experiment. Only two participants in the conditions with neutral facial expressions and neutral postural cues judged the active agent's emotional state as neutral. Two participants in the condition thought that active agent was happy, while two participants judged the active agent as being angry. One participant in the Neutral condition judged the active agent as being sad (Figure 5.14). The passive agent in the Neutral conditions was judged as afraid by three participants, neutral by two participants, surprised by one participant and sad by the last participant.

The participants' tendency to judge the agents' emotional states as affective even when the behavioural cues displayed were designed to be non-affective continued in their judgement of the agents'



(a) Angry Conditions (b) Sad Conditions

Figure 5.13: Participants' judgement of passive agent's emotional state towards the active agent.

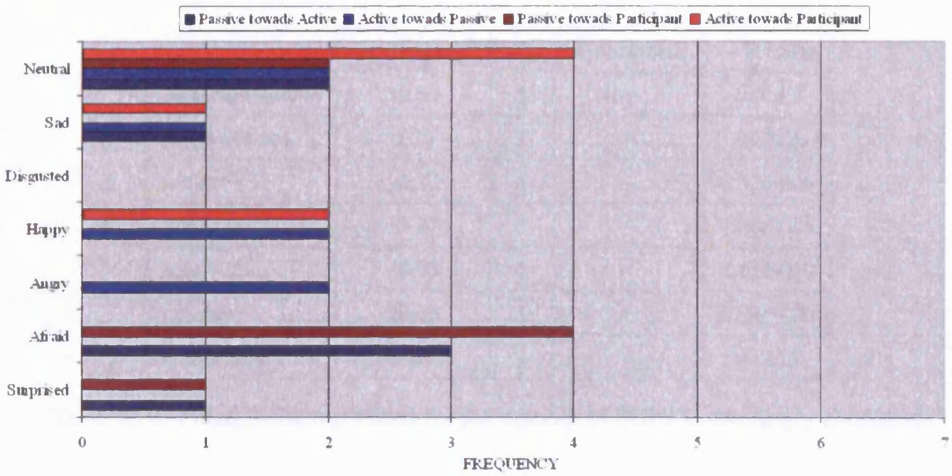


Figure 5.14: Agents' perceived emotional state towards each other in the Neutral condition.

emotional state towards the participant. The agents were designed to display only conversational feedback behaviours during the periods in which the participant interacted with the agents in a conversation. Yet only 27 out of 49 participants accurately judged the emotional state of the active agent as neutral. This increased to 33 out of 47 participants in the case of the passive agent. The second most common emotional state attributed to the active agent was surprise. In the case of the passive agent, six participants attributed an emotional state of surprise to the agent while six others attributed an emotional state of afraid. Further details of the participants' judgement of the agents' emotional states towards participants is available in Appendix F.6.

5.7.2 Physiological data

Even though three types of physiological recordings were collected during the experiments, only the Galvanic Skin Response (GSR) data was analysed. The physiological data set from the experiments were categorised in accordance to the condition of the experiment experienced by the participants. Each subset of data was then evaluated in aggregate.

5.7.2.1 Galvanic Skin Response

Further analysis was conducted on the number of skin conductance responses (SCRs) in the participants' GSR recordings as the main variable (Section 3.4.2.1). The number of SCRs was calculated for the whole period in which participants *approached* and *interacted* with the agents. There were three sets of SCR data for each participant corresponding to the three times the participant interacted with the agent. Therefore, there were three factors to take into account in the linear regression analysis: the type of postural cue, the type of facial expression and the number of time the participant had met the agents previously (Visit). The analysis focused on uncovering factor-related differences by fitting the same variables mentioned in Section 5.7.1. The log-linear model included the rate of SCRs recorded during the baseline period of the experiment and the duration of the experience in order to eliminate effects due to those factors.

	Deviance χ^2	**d.f.**	**Association**	**$\sim$ P value**
Neutral Posture	10.63	1	+	0.0011
Neutral Face	9.25	1	+	0.0023
APT	20.39	1	−	6.32e-006
VR	19.37	1	−	1.08e-005
Age	40.33	1	−	2.15e-010
Literate	20.46	1	+	6.09e-006
Gametime	4.63	1	−	0.0315

Table 5.10: Fitted linear regression for the number of SCRs in the Angry conditions

In the Angry conditions, both neutral postures and neutral facial expressions were positively associated with the number of SCRs (Table 5.10). The participant's level of computer literacy (Literate) is also positively associated with the number of SCRs. The other significant variables were all negatively associated with the number of SCRs and included the participant's age, their experience with VR systems, Gametime and the participant's measure of affect perception (APT). The overall model fitted had a deviance (χ^2) of 146.4 on 47 d.f. At the 5% level of significance, χ^2 has to be less than 64.0 at 47 d.f. for a model with good fit, so the overall explanatory power of the model is low.

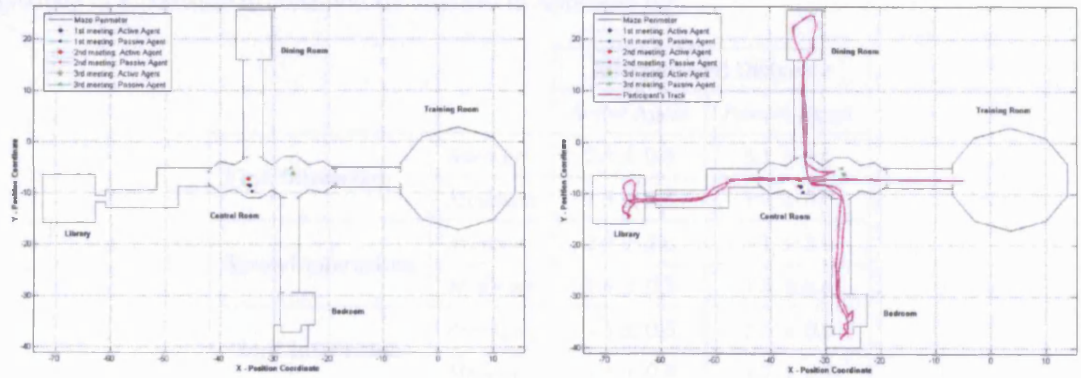
In the Sad conditions, the number of times (Visit) the participants had interacted with the agents previously was a significant factor in the fitted model. The second period during which the participants approaches and interacts with the agents is positively associated with the number of SCRs while the third period is negatively associated with the number of SCRs. Like the Angry conditions, neutral facial expressions were positively associated with the number of SCRs, however, unlike the Angry conditions, neutral postural cues were negatively associated with the number of SCRs (Table 5.11). The variable Gametime is also positively associated with the number of SCRs while the participant's measure of emotion contagion, their experience with VR systems and their age are all negatively associated with the

	Deviance χ^2	d.f.	Association	$\sim$ P value
Neutral Posture	3.95	1	$-$	0.0468
Neutral Face	7.59	1	$+$	0.0059
Visit	7.59	2		0.0044
Visit(2)			$+$	
Visit(3)			$-$	
EC	27.05	1	$-$	1.98e-007
VR	10.60	1	$-$	0.0011
Age	4.91	1	$-$	0.0267
Gametime	5.03	1	$+$	0.0250

Table 5.11: Fitted linear regression for the number of SCRs in the Sad conditions

number of SCRs. In addition to the rate of SCRs in the baseline period and the duration of the period of interest, the overall model fitted included the type of postural cue, the type of facial expression, the participant's age, their experience with VR systems, their measure of emotion contagion and Gametime. The model had a deviance (χ^2) of 79.9 on 46 d.f. and the significant explanatory variables had the deviances shown in Table 5.11. At the 5% level of significance, χ^2 has to be less than 62.8 at 46 d.f. for a model with good fit, so the overall explanatory power of this model is low as well.

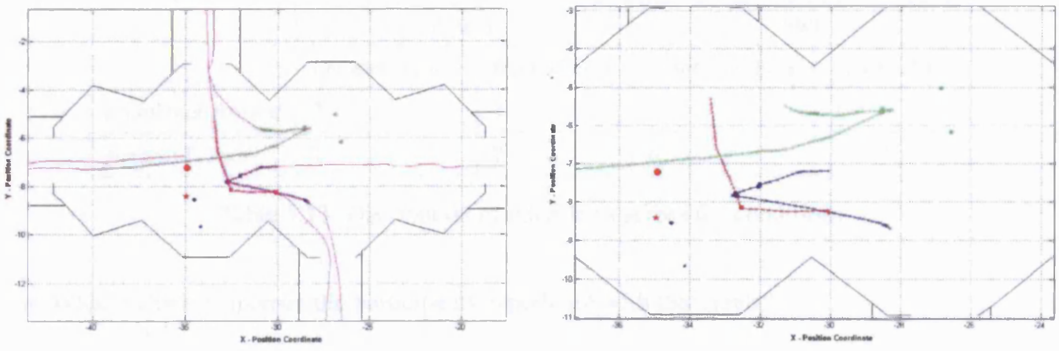
5.7.3 Tracking data and events



(a) Map of the virtual maze (b) One participant's visit

Figure 5.15: Map of the virtual maze and tracking data of a participant. The coloured circles show the positions of the active agent while the coloured stars show the positions of the passive agent.

During the analysis of the participant's spatial behaviour in the virtual maze, particular attention was paid to the durations in which the participants interacted with the agents. Each participant interacted with the agents three times, therefore, there were three sets of data points for each participant. Figure 5.15 shows a map of the maze along with a sample of the tracking data extracted from the logs of one participant during the three participant-agent interactions. The mean and minimum distances maintained by the participants to each agent was computed after extracting the relevant data from the tracking logs as discussed in Section 3.4.3.



(a) Tracking focused in the central room (b) Relevant tracking data in the central room

Figure 5.16: Tracking data of a participant in the central room.

Without exception all participants started the participant-agent interaction while they were stationed in front of the agents. Most participants continued to stand in the same spot throughout the interaction as well. This can be observed in the visualisations of the tracking data in Appendix F.7. Independent of condition, during the first interaction, on average, 44 out of the 49 participants maintained a smaller interpersonal distance with the active agent in comparison with the passive agent. The number of participants who maintained a smaller interpersonal distance with the active agent reduced to 41 during the second interaction and to 31 in the final interaction. However, an analysis of variances (ANOVA) between conditions did not reveal any significant differences. More visualisations of the tracking data grouped in accordance to condition are attached in Appendix F.7

		Interpersonal Distances	
		Active Agent	*Passive Agent*
First Interaction	*Average*	2.8 ± 0.6	3.1 ± 0.6
	Minimum	1.3 ± 0.8	1.8 ± 0.6
Second Interaction	*Average*	2.9 ± 3.0	3.1 ± 3.0
	Minimum	1.4 ± 0.5	1.8 ± 0.6
Third Interaction	*Average*	3.5 ± 0.6	3.5 ± 0.6
	Minimum	1.6 ± 0.8	1.7 ± 0.7

Table 5.12: Means and standard deviations of average and minimum interpersonal distances maintained by the participants during the three participant-agent interactions in the virtual maze.

5.7.4 Findings from Interviews

Each participant was interviewed on all the areas outlined in Section 5.6.1.5 and Appendices E.9 & E.10. Out of 49 participants, 2 were excluded since the recordings of the interviews were of poor audio quality. The remaining 47 interviews were distributed across the conditions (Table 5.13). Each audio recording of the interview was transcribed and analysed using the method (Section 3.4.4).

The following research questions were covered while analysing the interviews:

- What were participants' impressions of the agents?

	Angry		Sad	
	Emotional Face	*Neutral Face*	*Emotional Face*	*Neutral Face*
Emotional Posture	7	7	6	7
Neutral Posture	7	7	6	-

Table 5.13: Distribution of interviews across the conditions

- What realistic responses did participants experience with the agents?
- What were the participant's impressions of the agents' behaviour?

Out of forty-nine participants, eleven reported that their behaviour and responses in the maze were not realistic. For instance, some participants did not feel the need to maintain social norms in the presence of the agents since the agent did not look real. Similarly, these participants did not feel intrusive while they were approaching seemingly busy agents. On the other hand, some participants did not feel intrusive since they had a valid excuse in that they were given the task of asking the agents to guide them through the virtual maze.

About half the number of participants (23) reported that their responses and behaviour in the virtual maze was surprisingly realistic while thirteen participants reported experiencing a mixture of both realistic and unrealistic responses during the experiment. During the analysis of the semi-structured interviews, it became clear that the interpretation of behavioural cues in immersive virtual environment depended heavily on the character of the participant (interpreter). Some participants interpreted the behaviour of the agents in unexpected ways. For instance, one participant reported that the active agent was surprised at the detailed answer the participant had given during one of the participant-agent interactions in the experiment: "*I was asked to say what was in the room, the expression to what I had said was different to the 1st time and the guy looked a bit surprised.* (P7112)". This comment is especially interesting as the agents were designed to display no affective cues during the participant-agent interaction periods of the experiment. All responses triggered by the experiment controller followed a rigourously maintained pattern and yet the collection of participant responses were far from homogeneous.

The analysis focused on the variety of realistic responses and interpretations of scenario reported by all the participants. Out of the thirty-six participants who reported experiencing realistic responses, two participants attributed their responses to the humanoid form of the agents. However, most participants (34) reported that their responses were elicited through the behaviour of the agents: "*I was thinking thank God, this is not a real environment because you know, the way they were going at each other, I would probably rather not approach them at the moment. So that seemed quite; how do you say? Realistic.* (P3121)". One participant drew on his previous experiences and compared his responses to the agents in the virtual maze, to similar situations in the physical world: "*I think I guess it was fairly realistic, yeah because I remember one of the characters even saying 'pardon me', or something like, maybe he didn't actually hear what I said or maybe he recognised? that was surprisingly real because I am foreigner, people don't sometimes get what I say.* (P4211)".

5.7.4.1 Impressions of the interaction between the agents

Regardless of experiment condition, participants reported that their judgements of the interaction between the agents were made based on two cues: the body language of the agents and the audio properties of the verbal interaction between the agents. Only two participants reported using the agents' facial expression as a cue to the agents' underlying emotional state. One of the two participants reported using the passive agent's facial expression even though there were no facial expressions associated with the passive agent in the experiment. Out of forty-seven participants, sixteen participants reported using the agents' postural cues (body language) to deduce the agents' emotional states. Five participants reported using the agents' overall behaviour to judge the agents' emotional state. Eleven participants reported using vocal cues while ten participants reported using both body language, overall behaviour and vocal cues to judge the agents' emotional states. Three participants did not remember much about the agents since they were very task oriented while two participants viewed the agents as virtual objects and were unable to assign a psychological state to them. Six participants reported that, although they used the body language and/or the overall behaviour of the agents to interpret the underlying emotional state of the agents, the behaviour of the agents did not seem realistic.

	Angry		**Sad**	
	Emotional Face	*Neutral Face*	*Emotional Face*	*Neutral Face*
Emotional Posture	7 Arguments	4 Arguments		1 Argument
		2 Talking	3 Talking	4 Talking
			2 One-sided	2 One-sided
		1 Complaining	1 Complaining	
Neutral Posture	2 Arguments		1 Argument	-
	2 Talking	5 Talking	4 Talking	-
	3 One-sided	2 One-sided	1 One-sided	-

Table 5.14: Distribution of participant interpretations across the conditions

Generally, participants reported the interaction between the agents as either a conversation/discussion or an argument/disagreement. In either case, participants interpreted the behaviours of the agents in a holistic manner. Nine participants reported that the agents looked engrossed in their conversation since the agents only responded to the participants after the participants were within a certain distance: "*They seem to be busy with themselves because they didn't even like look at me or anything when I walked into that room. They obviously didn't notice that I was there until I was literally right in front of them.* (P3121)". Out of the fifteen participants who reported that the agents were involved in an argument, eleven were in the conditions where the active agent displayed angry postural cues towards the passive agent and two were in the conditions where the active agent displayed only angry facial cues towards the passive agent (Table 5.14). The eleven participants used body language as the main cue for their interpretation: "*His (active agent) body language was a bit aggressive. Not with me but with the other guy (passive agent). You can see him (active agent) bending and shaking his hands, his arms actually and he was bending, he must have been desperate. The other guy (passive agent) was just*

assuming all the blame, just nodding his head. Because you don't really know what they were talking or arguing about but the attitude was one (passive agent) of submission and the other one (active agent) of aggression. (P2111r)". Other terms participants used to describe the interaction between the agents were talking (20), a heated one-sided conversation where the active agent was verbally chastising or disagreeing with the passive agent (10) and or complaining to the passive agent (2).

5.7.4.2 Awareness and Reactive Behaviour

A secondary issue dealt with in one of the post-experiment questionnaire (Appendix E.5), partially to check if the agent behaviours were portrayed appropriately, was the responsiveness of the agents. Participants were asked to indicate if they thought each agent were a) responsive to them and b) interacted with them. Figure 5.17 shows that all the participants felt that the active agent was responsive to them and interacted with them. Interestingly a few participants felt that the passive agent was responsive and even more surprisingly three participants felt that the passive agent was interactive.

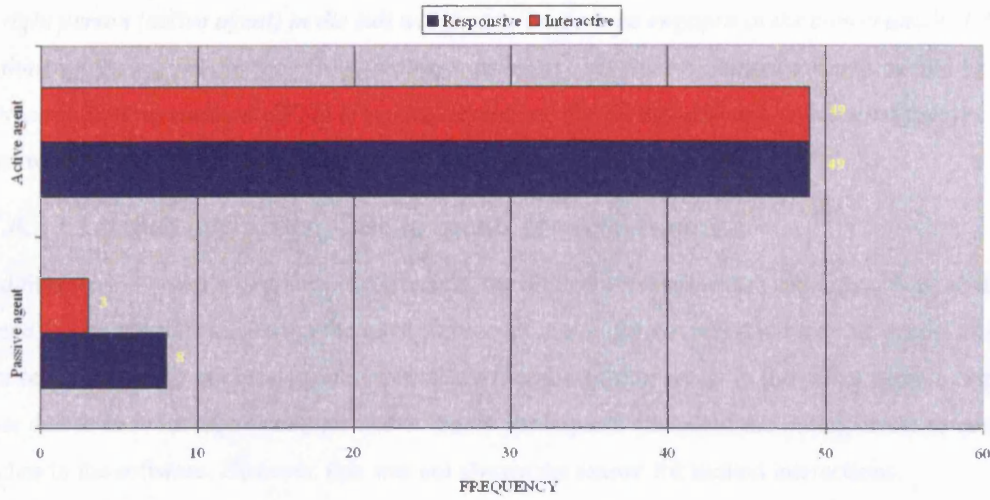


Figure 5.17: Participant's judgement of the responsiveness and interactivity of the agents

Participants reported that the agents were aware of the participant's presence and responded appropriately. This perception was mainly due to the participant's interpretation of the agents' behaviour but also in certain circumstances when the active agent appeared to have knowledge of either the participant's experience or the spatial properties of the virtual maze (Figure F.7). The participant's response to the questionnaire item was used during the interview process to uncover the aspects of the agents' behaviour that led to the participant's perception that the active agent was more responsive than the passive agent.

Forty-four participants reported that they felt that the agents were aware of them. The participants reported that they got this impression because the agents noticed the participant (8), acknowledged the participant's presence (5), halted the conversation that was already going on between the agents (21), turned towards the participants (41) and maintained eye-contact (33). One of the most significant behavioural cue that participants reported as being characteristic of perceived awareness was gaze (eye contact). Perceived awareness was associated with the agents's perceived ability to notice the participant's presence by halting their conversation, turning towards the participant and responding to the participant's queries while maintaining eye contact: "*Purely by when you walked to them and they turned around, both facing you and in a way, almost introduce themselves or ask 'how can we help'. That definitely was a sense that 'ah, ok I am now in the presence of these two people'.* (P7112)". Fourteen participants also reported that the agents had a sense of awareness when the agents appeared to have some knowledge pertaining to either the spatial properties of the virtual maze or the participant's experience in the virtual maze: "*I would say especially after the 3rd time, when I was looking for the door and I was told to turn behind me, so I assume that they must have been aware of my presence because otherwise it would be difficult for them to know.* (P5211)".

This effect was especially pronounced in the case of the active agent. Participants were more certain about the perceived awareness of the active agent either because they paid more attention to the more interactive active agent or because the passive agent did not engage them in an active interaction: "*I think the right person (active agent) in the suit was aware, because he engaged in the conversation. I don't, I am thinking the left person (passive agent) was aware as well since he turned towards me but he didn't make any effort to converse.* (P7111)". As expected, nearly all the participants reported that the active agent was more responsive and therefore focussed mainly on the active agent.

5.7.4.3 Limited interaction: Due to agents or social context?

In addition to the visual appearance of the agents, the limited vocal responses and interaction capabilities offered by the agents was cited as the most significant reason for not responding to the agents as if they were real. Thirteen participants were especially attuned to timing errors in the active agent's responses either due to network lags or human errors. These participants attributed the timing errors to computer glitches in the software. However, this was not always the reason for limited interactions.

The level of interaction between the participant and the agent depended on the participant's expectations. Thirteen of the forty-seven participants did not expect any queries or complex responses from the agents. A further twelve participants reported that they felt they were in the virtual maze to complete the task. Therefore, these participants did not attempt to engage the agents in conversation since they felt it would be fruitless or unnecessary: "*Because of the task and you don't expect the virtual people to have all the responses available for you, you just think they are going to say go to the room and that is all you think you will get, so I definitely didn't do much more than just ask for directions or something.* (P7112)". Since the task involved using the agents as guides to explore the virtual maze, the participants paid very little conscious attention to the agents. However, in cases where the participant did not interact with the agents due to the nature of the task, the participant did not report unrealistic responses to the agents.

A few participants did not attempt to engage the agents in an interaction initially because they had no expectations of the agent's abilities. These participants grew more comfortable with the interactions after they had gauged the agent's abilities through experience. When the agents respond in a manner that surpassed the participant's expectations, participants attributed a sense of realism to the agents:

"They caught me off guard a couple of times because they asked me what I had seen in the room, and I didn't anticipate them in doing that. So that brought an extra bit of realism to them. (P1211)". Five participants were surprised to experience responses to the agents especially when the participant had a preset expectation of the agents' ability: *"One thing is exactly like I expected which is like you know graphically they just won't look like real people, and therefore, you will find it hard to take them seriously as real people. But the fact that they were talking and seem a bit negative... that's a bit weird... was a real sort of bad vibes going on there. Which you know bad vibes from a bunch of graphics. That was maybe a bit more than I expected.* (P3121)". However, four participants expected the agents to be highly informed, and have better behavioural capabilities, despite their visual appearance. This impression might have been developed after the participant was told that the agents would act as guides to the virtual maze. For instance, one participant expected to conduct a more naturalistic and interactive conversation with the agents: *"It was a more like a automatic relation, you know. I ask for information, he provided the information. It was a more descriptive relation, ok, you have to do that, and I was waiting for the information to continue.* (P2121)". After unsuccessfully trying to engage the agents in a more complex interaction, these participants resigned to simply getting the information required to complete their task.

Some participants reasoned that the limited interaction was either due to the social context within which the interaction took place, a lack of the agent's interest in the participant or the agent's previous emotional state. Four participants attributed the uneasy interaction to the emotional state of the agents during the approach periods: *"Regarding the character on the left (passive), its because he seemed completely tied up with his own problem and didn't really turn to me and didn't try to make, try to really communicate with me at all, and with the character on the right (active), I think its, he seemed dutiful rather than welcoming. More to do in what wasn't said, rather than what was said. I think normally if someone was actually being friendly, they would ask you out of politeness or courtesy, what you thought of the experience or what, lets say what you found out in the rooms.* (P4121r)". Six participants attributed the limited interaction to the agent's lack of interest in the participants: *"I got more of a response out of the right guy (active agent), he seemed like he was you know quite confident. They both seemed pretty disinterested. Pretty kind of, you know, focused on themselves.* (P6212)".

5.7.4.4 Participant responses to the agents

One of the most significant responses reported by participants was the feeling of intrusion when interrupting the agents at the end of the approach periods of the experiment. This was especially the case when the agents were perceived to be involved in an argument or disagreement with each other. Twelve participants reported feeling hesitant, embarrassed and uncomfortable upon intruding on the agents' conversation: *"I walked up to the 2 people that seemed to have been in an argument. And since I didn't want to interfere with the argument, I just asked them that I had a task and I needed help. The main thing was that I saw them in a conflict, some type of argument, so I held back interacting with them. Similarly if I saw two people fighting in the real world, I wont approach the person anymore... I would choose the one that weren't in conflict with each other... I kind of felt bad for the person on the left (passive agent).* (P7111)". In addition to feeling intrusive, one participant also reported feeling anxious: *"I didn't really* *want to get to know him because unknown persons makes me more anxious so... and I am already not very good understanding conversation* (P3112)".

Another response that surprised nine participants was their tendency to be automatically polite towards the agents despite knowing that the agents were not real: "*I didn't think they would be able to have a discussion back with me. So I anticipated a very limited capacity because they were generated by the computer, so I just sort of spoke to them 'hello, which room am I suppose to go into?'. I was aware that I was polite though because I kept saying 'thank you' after they told me where to go which obviously is unnecessary because they are not real.* (P1211)". Some participants reasoned that this was because the agents were more realistic than they had expected. One participant reported feeling hesitant in carrying out his plan to try and listen in on the agents' conversation since it would have been socially impolite: "*I was kind of curious to look at them but then I probably moved because it felt socially awkward to just... because you did have the sense that they were responding to you... It felt like if I just kept standing there and staring at them, it would have been kind of inappropriate.* (P6212)". This is particularly interesting since the participant's curiosity to test the agents was overcome by his desire to maintain social norms.

5.7.4.5 Participants' impressions of the agents' persona

The agents were designed to have a neutral attitude towards the participants, however, the qualitative analysis of the interviews uncovered that participants had varying impressions of the agents' attitude towards them. Over 45% of the participants reported that the agents had a formal or neutral attitude towards them with slight variations. In keeping with their design, the agents' behaviour was described as polite, slightly upset at being interrupted but still neutral, friendly but cold and formal, sufficiently accommodating, instructional, in keeping with the situation, dutiful and generally behaved as can be expected.

Other participants reported different attitudes for each agent since the active agent was more interactive than the passive agent. Even though the only difference between the behavioural cues displayed by the agents towards the participants was in the level of subtlety, the overall effect seems to have created a noticeable difference between the apparent personalities of the agents. Participants interpreted the behaviours of the agents in a holistic manner by comparing the active agent with the passive agent in parallel: "*I thought he (active agent) was kind of a warm person in a computer way and the other person (passive agent) was very cold. He just didn't actually interact at all.* (P5212)". Furthermore, four participants perceived a change in the agent's attitude even though there was none.

Thirty-nine out of forty-seven participants compared the interaction they observed between the agents with similar interactions they had witnessed in real life and assigned roles with higher authority to the active agent due to a combination of the perceived persona of the agents: "*He (active agent) was obviously the one dominating between the two of them and he looked more, I suppose authoritative because he had the jacket and smart* (P5111)"; and the attire of both agents: "*The kind of impression I got was of the guy in the suit (active agent) being the boss of the guy in the white t-shirt (passive agent) basically. He worked for him or something. Maybe that was to do with the suit.* (P7221)".

Participants also assigned a position of higher authority to the active agent due to the manner in which they interpreted the behaviour of the agents: *"I think its probably the fact that there was no negotiation going on, I think friends tend to negotiate things even if they get angry, they tend to settle a negotiation. Whereas here, the role was one of the dominant and submission.* (P4121r)". Some of these participants went further than simply attributing the active agent with a higher status. The participants reported using their perception of the active agent's to choose the active agent for interaction since the active agent looked commanding, more engaging and looked more approachable. Furthermore, three participants reported playing a passive role in their interactions with the active agent because they felt intimidated and negatively judged: *"I think the guy on the right (active agent) looked at me but I think he remained quite distant and I had a sense he was kind of looking down to me.* (P4121r)". One participant felt that the active agent was unimpressed with their performance in the maze.

Participants attributed the passive agent's lack of responsiveness to non-technical constraints. This was especially the case when participants tried to interact with the passive agent and were unsuccessful. Participants reported that the passive agent was less friendly towards the participants due to the lack of interaction with the participants. Three participants felt that the passive agent was non-responsive because of earlier problems with the active agent. Twenty-six participants reported not paying attention to the passive agent. This effect was observed throughout to such an extent that participants could not accurately described the visual appearance of the passive agent. One participant did not even remember the gender of the passive agent:*"The person on the right hand (active agent) tells me what I had to know and they go back to their little thing and the person on the left (passive agent) might have said something, you know as in 'hi' or something like when I approached them but I don't think he or she gave me any directions at all.* (P3121)". Few of the twenty-six participants reasoned that the passive agent could not respond actively to the participant due to technical constraints. Six participants did not interact with the passive agent because the lack of responsiveness made the passive agent appear indifferent and cold. However, twelve participants did not pay attention to the passive agent because the active agent took control of the interaction and appeared to be the agent in charge before the participant's interaction period with the agents. Yet, the passive agent's behaviour was not judge as unrealistic. It was generally accepted that the behaviours portrayed by the passive agent for the role it played was natural and in keeping with the scenario: *"It was always the one on the right (active agent) which was addressing me and the one on the left (passive agent) never said anything. So I guess, in that way I was responding as if they were real people because I expect if two people were having an argument and one was dominant, and then you came up to him and asked, you know for directions or something, it is going to be the one who is dominant in the argument that is going to say something.* (P7221)".

In keeping with participants' feelings of intrusion (Section 5.7.4.4), two participants perceived the agents as being unfriendly and described the agents' behaviour as unhelpful and dismissive. Five other participants perceived the agents as not welcoming, impatient and too engrossed in the previous interaction with each other. Since the active agent was more heavily involved in the interaction between the participant and agents, the active agent in particular was perceived as impatient. One participant reported that the agents were trying to cut the interaction short since the agents were interested in continuing their previous conversation with each other: "*They were trying to get rid of me because they wanted to talk about... what they were talking about... So one guy (active agent) was instructing or telling the other guy (passive agent) off. And he was silent. It wasn't his place to talk to me. Because the other guy was taking control of the conversation and continued the conversation whenever I came.* (P2112)". Another participant reported that the agents did not want to have an interaction with him initially but this changed during the experiment. More of the comments from participant interviews are available in Section F.5 of Appendix F.

5.8 Discussion

This thesis follows the methodological standpoint presented by Slater (2004). Questionnaires by themselves are not sufficient for the interpretation of participant responses to virtual stimuli, but are a useful adjunct to a range of other tools including objective and behavioural measures. As such, a combination of both quantitative and qualitative methods was used to analyse the wealth of data collected in the virtual maze. The analysis was mainly conducted after splitting the participant data into two separate groups (Angry and Sad). The exception was in the analysis of the interview data.

The analysis indicated that the postures played a vital role in the way participants responded to the agents. In the Angry conditions, neutral postures were positively associated with reported copresence. This was mirrored in the Sad conditions. Additionally, neutral facial expressions were also positively associated with reported copresence in the Sad conditions. Neutral postures were positively associated with physiological arousal in the Angry conditions while in the Sad conditions, neutral postures were negatively associated with physiological arousal. Neutral facial expressions were positively associated with physiological arousal in both the Angry and the Sad conditions. This is surprising since not even one participant in the whole experiment reported using facial expressions to judge the underlying emotional state of the agent independent of whether the participant accurately recognised the emotional state of the agent. On the other hand, the importance of postural cues in the communication of affect is re-enforced by the results obtained from both the physiological responses of participants and subjective questionnaire responses. However, in this case, it appears that the Neutral postures have outperformed the affective postures designed to portray Anger and Sadness. The analysis of the interviews suggest that the affective postures played a crucial role in preventing participants from projecting their perceptions about the agents onto the scenario. On the other, the neutral postures seem to allow participants to interpret the scenario as they saw fit.

In the Sad conditions, subjective results, as collected through a questionnaire to judge the agent's emotional state and participant interviews, suggest that participants were not able to accurately recognise the active agent's emotional state as Sad. This suggests that the parametric model of affective postures, for the Sad conditions, did not accurately portray cues associated with Sadness. This effect is especially pronounced in the Sad conditions and reflected by one participant's comments: "*I thought they seemed to be engrossed in some kind of conversation, although the guy kind of looked awkward. He was sort of bent down and one leg stuck out and was bobbling about.*". This partially explains why the neutral postures and neutral facial expressions outperformed the affective cues in the Sad conditions. One reason for the effects experienced may be in the underlying posture model (Coulson, 2004) used in the conditions. In addition, eight out of the fourteen participants, in the Sad conditions with affective postures, attributed an emotional state of Neutral to the active agent, while a further four participants thought the active agent was Angry. Considering Paterson et al.'s (2001) results, which suggested that body movement played a key role in the portrayal of Sadness, it is possible that the lack of cues in the quality of body movement prevented participants from accurately recognising the postures. In addition, it is possible that the qualities of the verbal conversation between the agents during the approach periods were not in keeping with a Sad scenario. Although participants could not understand the contents of the conversation, the quality of the conversation was essentially neutral and emotionless. It is possible that participants found it easier to attribute Angry connotations to a Neutral conversation in comparison to Sad ones. It is also possible that the muffled conversation played a role in allowing participants to project their perceptions onto the agents when the agents displayed neutral postural cues.

The results obtained in the Angry conditions were slightly more puzzling. Generally, a higher physiological arousal is expected to be associated with a higher reported presence/copresence. This association is apparent in the Angry conditions. Therefore, it is clear that the Neutral postures did outperform the Angry postures with respect to eliciting realistic participant responses. Yet this is not because participants did not recognise the underlying emotional state of the active agent. Eleven out of the fourteen participants, in the Angry conditions with postural cues, accurately recognised the emotional state of the active agent as Anger and reported that body language was a primary indicator of the active agent's underlying emotional state. Furthermore, in the conditions with congruent angry behavioural cues, all seven participants accurately recognised the underlying emotional state of the active agent. This is in keeping with Montepare et al.'s (1999) argument that postures and body movement are more important for judging an individual's psychological state especially if the individual's facial expressions are not visible. The results suggest that even though the parametric model for affective postures used in the Angry conditions portrayed postural cues associated with Anger, the resulting behaviour was not good enough to elicit appropriate physiological responses.

Surprisingly the agent's behavioural fidelity did not seem to have had an effect on the social distances maintained by the participants with the agents. For instance, it was reasonable to expect that a participant approaching an Angry agent might be fearful or apprehensive and therefore maintain larger interpersonal distances with the agent while a participant approaching a Sad agent might be sympathetic and therefore maintain smaller interpersonal distances with the agent. However, no significant differences were detected in the analysis of the tracking data. This might be due to two reasons. It is possible that the participants' perception of the agents did not affect the interpersonal distances maintained by the participant because of the nature of the task. Another reason for the lack of factor-related differences could be due to the mechanism used to trigger the agents into turning to focus attention towards the participant. Independent of condition, the agents are triggered to stop their conversation and turn towards the participant at the same distance. This could in turn be the factor that controls the distance at which participants interact with the agent. It was also apparent that participants felt more comfortable in adjusting the participant-agent interpersonal distance during the second and third interaction periods rather than the first. This suggests that participants were acting on instinct during the first interaction but that the illusion broke down after repeated exposure to the same agents. Further experiments are required to test these theories.

Many participants remarked on the appearance of the agents by referring to the active agent as having authority over the passive agent. In the Angry conditions with affective postures, the participants continued to infer that the passive agent was being chastised due to some blunder on his part. They refer to the body language of the active agent as being aggressive and agitated towards the passive agent. These participants also referred to the conversation between the virtual agents as an 'argument'. Less than 30% of the participants in the Angry conditions with postural cues identified the passive agents as Neutral. The passive agent was perceived to be afraid by 50% of the participants in the condition even though the cues portrayed by the agent were always neutral. This suggests that participant's perception of the state of the active agent was influenced by their perception of the whole social relationship between the two agents. In other words, since participants perceived the situation as an argument, if one agent was angry then the other by implication should have been afraid. In the Neutral and Sad conditions, references to the body language of the agents indicating an argument was rarely made, however, the active agent was at times judged as being angry. In these instances, the passive agent was judged as either afraid or neutral. This suggests that it is equally important to design appropriate behavioural cues for the passive agent as it is for the active agent since the behaviour of the passive agent influences participant's interpretation of the active agent's emotional state. However, the paradigm set up to study the role of the body in communication affect, using two agents, seems to suit the purposes of the experiment.

Finally, a number of realistic responses were reported by the participants in the post-experimental interviews including feelings of intrusion and the desire to maintain acceptable social norms with the agents. The participant's tendency to maintain social norms with the agents was confirmed through the analysis of tracking data. This is in keeping with the operational approach adapted in this study. It suggests that to some extent participants in IVEs do respond to agents as if they were real even when the behavioural cues displayed by the agents are not completely accurate. This was verbalised by a participant in the Angry condition: "*For me the main thing is that I saw them in a conflict, some type of argument, so I held back interacting with them. Similarly if I saw two people fighting in the real world, I won't approach that person.*".

5.9 Summary

Previous research suggests that the postural cues displayed by individuals in the physical world play an important role in the communication of affect (de Gelder, 2006; Dittmann et al., 1965; Montepare et al., 1999). If the same theory is appropriate in IVEs then designing high-fidelity postural cues might be a cost-effective approach to designing effective full-body virtual humans for use in collaborative VEs. The experiment presented in this chapter was designed to investigate the role of postures in the communication of two affective states: *Anger* and *Sadness*. Sections 5.1 to 5.3 dealt with the design of

Virtual Human	Agents
Factors	Type of Postural cue Type of Facial expression
Participants	49 males
Environment	Virtual Maze
Apparatus	ReaCTor Physiological devices
Software	DIVE VRPN (C++ Plugin) PIAVCA (C++ Plugin) PIAVCA posture loading extension (C++ Plugin)
Data Collected	Questionnaires Interviews Physiological measures Proximal measures
Data Analysed	Questionnaires Interviews Physiological (GSR) measures Proximal measures
Publications	Vinayagamoorthy et al. (2006a)

Table 5.15: Summary of the experiment on posture

the experiment and parametric model of affective postures used while Sections 5.4 & 5.5 presented the population and procedures used in the experiment.

A wide variety of participant responses were measured (Section 5.6). The findings from these measures suggested that the parametric model of affective postures presented in Section 5.3.5.3 did not accurately portray Sadness since participants were unable to accurately recognise the emotional state of the agent in the Sad conditions. A significant number of participants accurately recognised the emotional state of the agent using Angry postural cues. In both conditions, the participant's reported copresence were positively associated with the Neutral postures. In the Angry conditions, the Neutral postures were also associated with higher physiological arousal. This suggests that although postural cues can portray the emotional state of Anger, the parametric model used in this experiment is incomplete.

The most interesting results came from the analysis of the participant interviews and their judgement of the emotion in particular their interpretations of the agents towards each other. Surprisingly, participants attributed a variety of emotional states to the passive agent regardless of condition. In cases, where the agents displayed neutral postural cues, participants attributed emotional states to the agents to fit their perceptions of what the agents were doing. This suggests that participants associated an emotional state to the Neutral postures based on the context in which it was displayed in. These findings are in keeping with those reported in Freeman et al. (2003), where participants attributed sentience to agents displaying essentially neutral cues. Furthermore, participants always associated the passive agent with an emotional state that complimented the emotional state of the active agent. In other words, the subjective assessment of the passive agent was dependent on the assessment of the active agent.It also supports the argument presented by Watzlawick et al. (1968) that all behaviours contain information and are therefore expressive.

Finally, tracking data which gave the position and orientation of the participant in the VE together with time stamps and event logs were also gathered during the experiment. The analysis of the tracking data did not reveal any factor-related differences (Section 5.7.3). However, participants maintained an acceptable interpersonal distance with the agents during the interaction periods of the experiment. This was in keeping with participant feedback in the post-experiential interviews, and presented further supportive evidence that participant responded, socially, to the agents (Section 5.7.4).

In summary, the results suggest that postures play an important role in the communication of affect but that the design of the parametric model generating the agents' postural cues is incomplete (Section 5.8). The findings suggest that it is better to design virtual humans with neutral postural cues in comparison to virtual humans with an incomplete or incorrect set of affective postural cues. Section 5.7 presented the results while Section 5.8 discussed the findings. Section F.8 in Appendix F gives a list of the main findings in this experiment. The next chapter presents an experiment designed to investigate the role of body movement in the communication of affect.

Chapter 6

Experiment: On Kinesics

Like the experiment discussed in Chapter 5, the experiment presented in this Chapter was designed to investigate the role of kinesics in the communication of affect in IVEs. This experiment extends the experiment discussed in Chapter 5, by exploring different qualities of body movement associated with the underlying affective state of agents using the same two emotions: *Angry* and *Sadness*. In order to eliminate the influence of facial expressions on participant responses, a face mask was used to obscure the agents' facial expressions.

The second most significant problem was the uncertainty faced when interpreting multiple types of participant responses to the agents. This was mainly due to the lack of definitive research dealing with the possible responses individuals experience during a given social situation. For instance, one of the questions investigated in the previous experiments was: *if a participant encounters an agent exhibiting behavioural cues indicative of anger, does the participant respond accordingly*? However, since each participant is likely to respond to the agents in a unique manner, it was difficult to judge if the participant would have reacted in the same manner had the agent been a real person. In this experiment, a novel methodological approach was taken to establish a baseline with which to analyse participant responses to agents. Each participant was exposed to two versions of an affective virtual agent (virtual conditions) in addition to a real person/professional actor (real condition) during the experiment. The measures which proved to be especially fruitful in the previous experiment were used to collect participant responses: physiological (GSR) and questionnaires.

The next section presents the hypothesis attached to this experiment while Section 6.2 discusses the justification for eliminating the use of facial expressions through the use of a mask. Sections 6.3 - 6.5 discuss the design and procedures of the experiment. Section 6.6 presents the participant responses measured in the experiment while Section 6.7 presents the analysis of the collected participant responses. Section 6.8 presents a discussion of the findings. Finally, Section 6.9 summarises this chapter.

6.1 Hypothesis

The previous experiment was designed to investigated the role of posture (and facial expressions) in the communication of an agent's emotional state. This experiment focused on improving the parametric model of posture used in the previous chapter and also investigate the role of body movement in the communication of an agent's emotional state. As with the previous experiment, the affective state of the agent/actor was focused at another agent as opposed to the participant, however, only bodily cues were used to portray the affect. The facial expressions of the agents were obscured through the use of a mask.

The main hypothesis was that a virtual human controlled by a parametric model of affective posture and body movement would play a significant effect on participant responses in IVE. The associated premises were that:

- Participants would respond to an affective agent, if the agent displayed appropriate affective bodily cues (postures and body movement).
- Participants would be able to accurately recognise the underlying emotional state of the agent, if the agent displayed the appropriate affective bodily cues.
- Participant responses to an agent portraying affective bodily expression would be similar to participant responses to an emotional actor in a similar situation.

The Null Hypothesis: There will be no differences found in the participant's responses (physiological and subjective) to the three conditions described in the following sections.

6.2 Justification for removing facial expressions

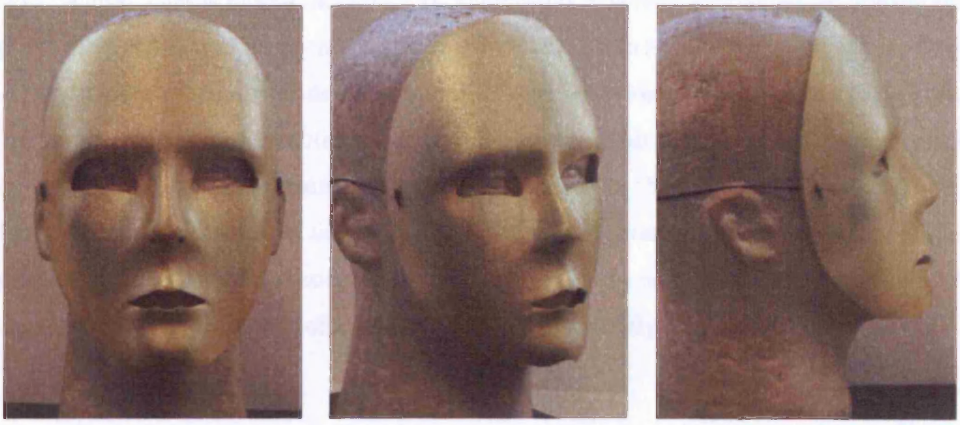
In the previous experiment, facial expressions had a significant role on participant responses. This was surprising since the virtual scenario, created for the experiment, involved placing the virtual agents in such a manner that encouraged participants to focus on the bodily cues portrayed by the virtual agents. It has been argued by James (1932) that one aspect of an expression colours the perception of the whole expression. In order to ensure that the participant responses recorded in this experiment were due to the kinesic cues displayed by the active character, it was important to eliminate any influence of facial expressions on participant responses.

In his paper, James (1932) argued that studying each component of a posture independently is justifiable for the purposes of analysis, but that all behavioural cues should be considered as a whole unit that function as a holistic expression of state in relation to the context. This is in keeping with the scenario adopted for this experiment. Additionally, in keeping with the argument for congruent behavioural cues made by Dittmann et al. (1965), de Gelder (2006), Montepare et al. (1999) and Planalp et al. (1996), the addition of a congruent facial component to a robust parametric model of affective kinesics can only enhance the emotion recognition rates and participant responses to the virtual human.

6.3 Experimental design: Building the scenario

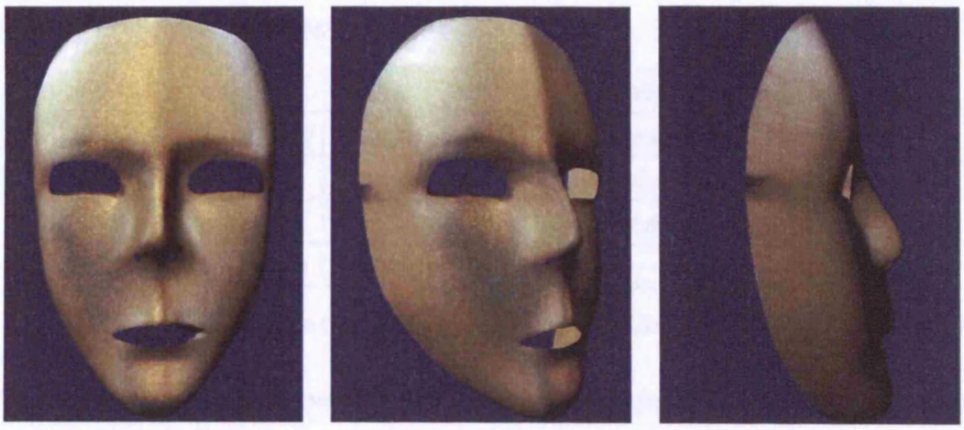
Like the previous experiment, the goals of this experiment involved measuring participants' automatic responses to the virtual agents. However, unlike the previous experiment, participants were exposed to the agents for a longer period of time. In this experiment, one of the goals was to build a more robust parametric model of affective bodily cues. Therefore, participants were asked to judge the quality of the agent's animation in comparison to the movement of a real person (a professional actor - (Gregory, 2006), Section G.1 in Appendix G).

Participants undertook the experiment in a virtual bar in which an active *character* (agent or real person/actor), playing the role of a barman, gave instructions to a passive agent. The active character was scripted to portray the appropriate affect (Anger or Sadness) through body movement. In order to completely eliminate any effects the perceived facial expression of the agents (or actor) would have on the participant responses, a virtual and real face mask was used to obscure the faces of the virtual agents and the real actor. A number of masks were considered and discarded due to the potential meanings attach to them. The mask, shown in Figures 6.1 and 6.2, was chosen, with the guidance of the professional actor, since it best suited the purposes of obscuring any facial expressions while remaining relatively bland and non-threatening.



(a) (b) (c)

Figure 6.1: The real mask.



(a) (b) (c)

Figure 6.2: The virtual mask.

In keeping with the terminology introduced in the previous chapter, in the two virtual conditions, one agent is referred to as the *active* agent while the other agent is referred to as the *passive* agent. In one of the virtual conditions, the active agent portrayed an affective state through the quality of pre-recorded audio and a combination of posture, gestures and body movement. This was referred to as the *moving* *virtual* condition. In the second virtual condition (the control condition), a static active agent portrayed an affective state through the quality of pre-recorded audio only. This was referred to as the *static virtual* condition. In the real condition, the active agent was replaced by the professional actor.

One of the significant decisions made during the design phase of the experiments was the use of a static virtual condition instead of a 'neutral' virtual condition. Initially the active agent in the control condition was designed to display a 'neutral' set of bodily cues along with pre-recorded affective audio cues, however, this design was abandoned for two reasons. Firstly the combination of a 'neutral' set of cues with a set of affective audio cues introduced incongruence into the agent's behaviour. Secondly, and more importantly, one of the interpretations from the findings of the previous experiment is that the idea of 'neutral' cues is flawed (Section 5.9). Participants generally interpreted the holistic scenario portrayed to them and judged the agents as affective even when both agents displayed supposedly neutral cues. A combination of randomly designed and displayed behavioural cues can lead participants into assigning various attitudes and emotional states to the agents. This was evident in the analysis of the interviews collected from participants in the previous experiment (Sections 5.7.1.5 and 5.7.4) and the second preliminary experiment (Sections 4.2.5.2 and 4.2.5.3). Although the agents were not designed to portray a specific affective state, participants judged agents using a number of different labels including cold, annoyed, arrogant, grumpy, indifferent, preoccupied, unfriendly, non-inclusive etc.

6.3.1 Design and variables

Two sets of within-group designs with three levels were used in the experiments: an Angry set and a Sad set. Each participant was exposed to one moving virtual condition, one static virtual condition and one real condition within the same affective set. In each condition, the participant observed two characters. In the two virtual conditions, the two characters were played by two agents. In the real condition, the two characters were played by one agent and one actor.

	Angry	**Sad**	**Characters**
Level 1	Static	Static	*An Active agent & a Passive agent*
Level 2	Moving	Moving	*An Active agent & a Passive agent*
Level 3	Real	Real	*An Actor & A Passive agent*
Number of Participants	13	10	-

Table 6.1: The two within-group designs for Angry and Sad

This design allowed the participant to compare and contrast within the conditions and therefore give a subjective evaluation at the end of the experiment. The moving virtual and the static virtual conditions were displayed, to the participants, as the first two conditions in random order while the real condition was always displayed last as condition 3. This order of presentation was maintained in order to avoid participants using the actor's portrayal as a template for judging the agents in the virtual conditions.

6.3.2 Designs, pilots and revisions

The pilots were used to design the scenario and task in the virtual bar, design an appropriate script for the active agent (or actor), generate behaviour for the active agent and acted as dress rehearsals.

Piloting the scenario and task: One of the key issues considered in the design of the experiment was the role of the participant. There were two main options. The first option was to assign the role of an observer to the participant. The participant would be asked to observe an active character and a passive agent having a conversation in a bar. Similar to the previous experiment, a passive agent would be the focus of the emotional active character's attention. The second option was to use only one agent which focused affective behavioural cues towards the participant (Figure 6.3). This option was discarded, after initial pilots, due to the differences in participant's perception of the mask. Even though the mask (Figures 6.1 and 6.2) was chosen very carefully, at some angles the mask seemed to present an illusionary smile or frown similar to the ones described by Mignault and Chaudhuri's (2003). This was especially pronounced in the IVE since the participant's height played a significant role in a direct participant-agent interaction. For instance, some participants looked down at the agent while others looked up at the agent. This, in itself, played a role in how participants interpreted the character's behaviour.

(a) Moving virtual condition (b) Static virtual condition (c) Real condition

Figure 6.3: An alternative scenario with the active agent or actor in front of the participants talking at the participant in an affective manner.

Once the first option was chosen, particular care had to be taken in the placement of the active character. This was to ensure that the active character's position was not outside the physical boundaries of the VR system (Trimension ReaCTor), since the real condition involved an actor swapping places with the active agent. The appropriate placement of the passive agent was less problematic since that issue was dealt with in the previous experiment. As part of the scenario, participants were asked to remain in the virtual bar for the duration of one song (2 minutes and 33 seconds). However, this duration varied depending on the length of the interaction between the passive agent and the active character. In the virtual conditions, the interaction between the passive and active agents depended solely on whether the conditions belonged to the Angry (2 minutes and 15 seconds) or Sad (2 minutes and 30 seconds) set. However, in the real condition, where the actor interacted with the passive agent, the duration varied slightly around the times stated above.

In early pilots, in order to keep the participants occupied and involved, participants were asked to inform the active character of an event. The event was defined as the change over of pictures in a virtual frame in the bar (Figures 6.4 and 6.5). However, this task encouraged participants to move around, and sometimes shout at, the agents when the agents were slow to respond. Since, participants were being physiologically monitored to detect event-related responses, any movement or physical activity on the

(a) (b) (c) (d)

Figure 6.4: The alternative task in the virtual condition.

(a) (b) (c) (d)

Figure 6.5: The alternative task in the real condition.

participant's part added noise to the physiological data. In the end, participants were simply asked to stand still and observe the two people having a conversation in the virtual bar. In order to elicit responses from the participant during the experiment, the agents were programmed to look at the participant about 5 times. Gaze behaviour and eye contact have proved to be an incredibly important social cue. In pilots, the act of looking at the participant, gave the participant the perception that the agents were aware of their presence.

Piloting an appropriate script: The creation of a context was an important part of designing this experiment. A professional actor (Gregory, 2006) was hired to play the part of the active agent in the real conditions (Appendix G.1). The actor needed a script and a back-story in order to portray an effective performance. A script was created based around the fictitious role of a barman explaining the different duties that need to be undertaken to close up the bar (Appendix G.2). A number of possible back-stories were created for each set of experiments (Appendix G.3). The following two back-stories were used in the final design of the experiment:

Angry:

You are the bar manager talking to a bar man. He's been here for weeks but never does his job properly. What really annoys you is that he doesn't clear up properly as you have to sort it out in the morning. This morning it was really bad, you had to work in a real rush to get everything looking OK before the bar opened, and you had to open half an hour late which meant turning customers away.

Sad:

You're a barman; you've been doing this job for years, and have had a heavy week and are sick of it. You are explaining to an employee who is also a mate all the annoying things you have to do every night.

Generating behaviour: The parameters to define the behaviour of the active agent was designed based on the portrayed behaviour of the actor. In this manner, participant responses to the agent and the actor could be compared and analysed to determine any similarities. The behaviour model used is discussed in Section 6.3.6.3. The actor went through several iterations of portraying the situation using the script and back-story in order to ensure that the same performance was given to every participant in the final experiment.

Dress rehearsals: Once the actor was satisfied with his performance, pilots were carried out to ensure that both the actor and the experimenter were familiar with the experiment procedures. These dress rehearsals were also used to make corrections to the questionnaires and information sheet used.

6.3.3 Apparatus

The same apparatus used in the previous experiment was used again with the exception of the navigation wand. The Trimension ReaCTor (Section 3.2.3.1) was used to display the VE while the ProComp+ device (Section 3.2.3.3) was used to record the participant's physiological responses. The participants behaviour and comments were recorded on video and audio tapes. The apparatus setup is illustrated in Figure D.1 in Section 3.2.3 of Chapter 3.

6.3.4 Software

Once again, the DIVE (Section 3.2.4.1) software with the VRPN (Section 3.2.4.2) and PIAVCA plugins (Section 3.2.4.3) were used to implement the experiment. The animation of the agents were controlled through the use of Tcl scripts and handled using the DIVE-Tcl interface with a Tk GUI. DIVE events were used to indicate the points in the log files at which the agents/actor looked at the participant.

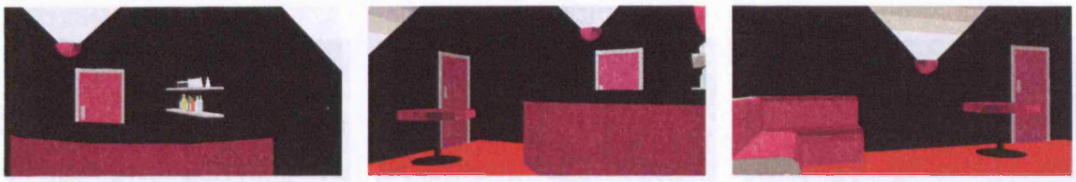
In addition to the core functionality of PIAVCA, a PIAVCA motion filter was developed to enable the modification of an existing animation file by changing it's properties over specifiable parts of the file. The motion filter was used to change the speed and amplitude of the mid-section of an animation file. The motion filter's input parameters were a *speed* factor, a *amplitude* factor, a *start delimiter* and an *end delimiter*. The start and end delimiters determined the part of the animation file to be altered. For instance, a start delimiter of 0.2 and an end delimiter of 0.1 indicated that the mid-section of the animation file was after the first 20% and before the last 10% of the file. The speed factor changed the speed of the animation over the section of the file defined by the start and end delimiters. The speed factor affected the animation by changing the temporal properties of the motion file. The effect of increasing the speed of an animation file over a large mid-section introduced an apparent *jerkiness* to the animation. The amplitude factor stretched the animation between the start and end delimiters. It changed the spatial properties of the animation file in the mid-section. This was used to introduce the *exaggeration* of certain gestures.

For instance, if the parameters input into the PIAVCA motion filter stated that the start delimiter was 0.2, the end delimiter was 0.1, the speed factor was 5 and the amplitude factor was 3. Then the first 20% of the original animation file would stay intact as would the final 10%. However, the middle 70% of the animation file would be sped up to 5 times its original speed and exaggerated to 3 times its original amplitude. This would make the resultant animation seem more energetic in the middle. As the speed of the animation is increased over shorter delimiters, the jerkiness of the animation at the beginning and the end of the animation becomes more pronounced since there is sudden reduction of speed over a small window.

6.3.5 The Virtual Environment

The virtual environment, used in the experiment, consisted of a virtual bar and a training room superimposed onto it. The training room was a simple room with three large single-digit numbers. Either the training room or the virtual bar was rendered visible at any one time. Unlike previous experiments, the training room only served to ensure that the VR system was working and that participant could see the VE displayed. This pseudo training process was conducted at the beginning of each experiment and in-between conditions.

6.3.5.1 The Bar

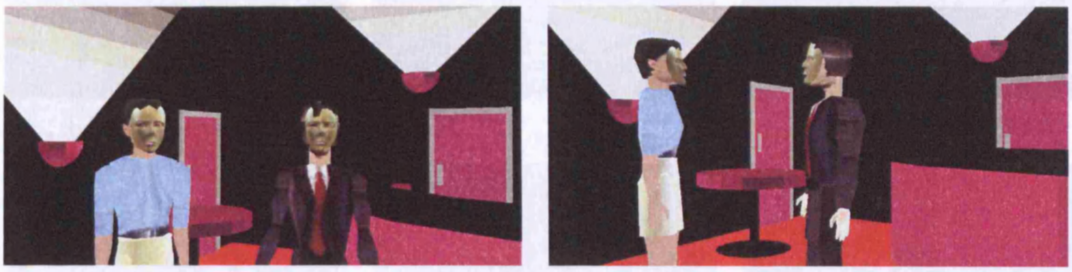


(a) Right (b) Centre (c) Left

Figure 6.6: The Virtual Bar

The virtual bar was a pre-existing model. In addition to bodily cues, the agents (and actor) were accompanied with affective audio cues. The audio cues were important to add realism to the scenario being portrayed. In order to mask the agent's audio cues, the virtual bar allowed for the playback of background music: "*Baby it's you*" by the Beatles.

6.3.6 The Agents: Inducing participant responses



(a) Mugshot (b) Actual placement

Figure 6.7: The masked virtual agents

The agents used in this experiment were similar to the ones used in the previous experiment (Section 5.3.5.1 of Chapter 5) and were assigned the same roles. In addition, face masks were used to occlude the facial expressions of the agents and actor in the experiment (Figure 6.7). This ensured that the only behavioural cues perceived by the participants were those portrayed through the body.

6.3.6.1 Overview of agent behaviour

Throughout the experiment, the agents were designed to carry out a conversation with each other in which the active character was the main speaker. In the virtual conditions, the active agent was designed to display condition-dependent affective cues (bodily and/or audio) while the passive agent was designed to display some 'neutral' cues. In the real condition, the active agent was replaced by the professional actor.

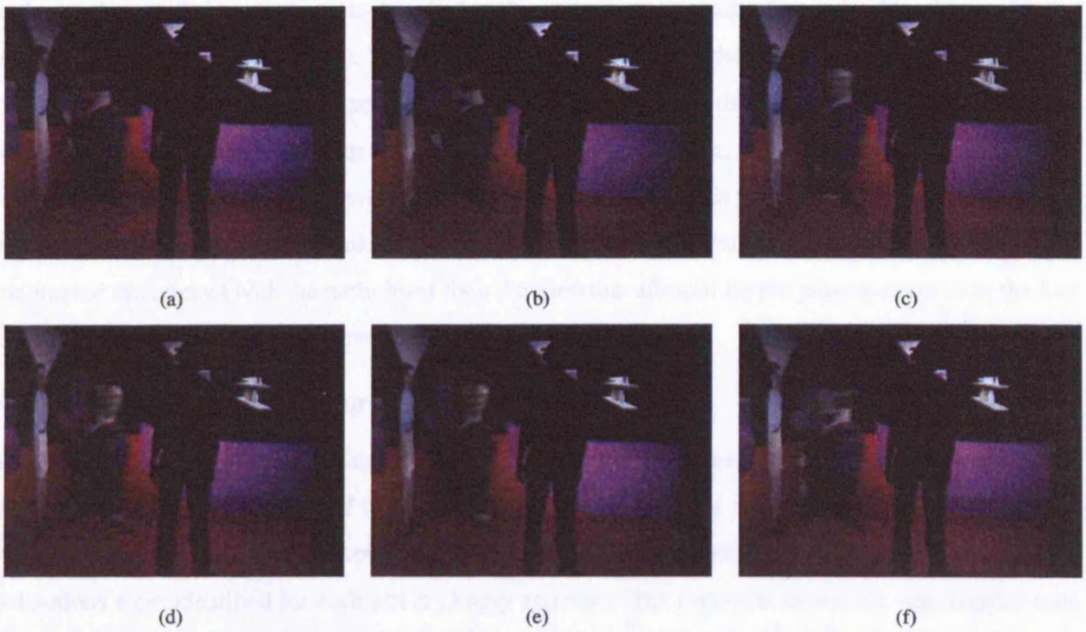


Figure 6.8: The turning process in the virtual condition

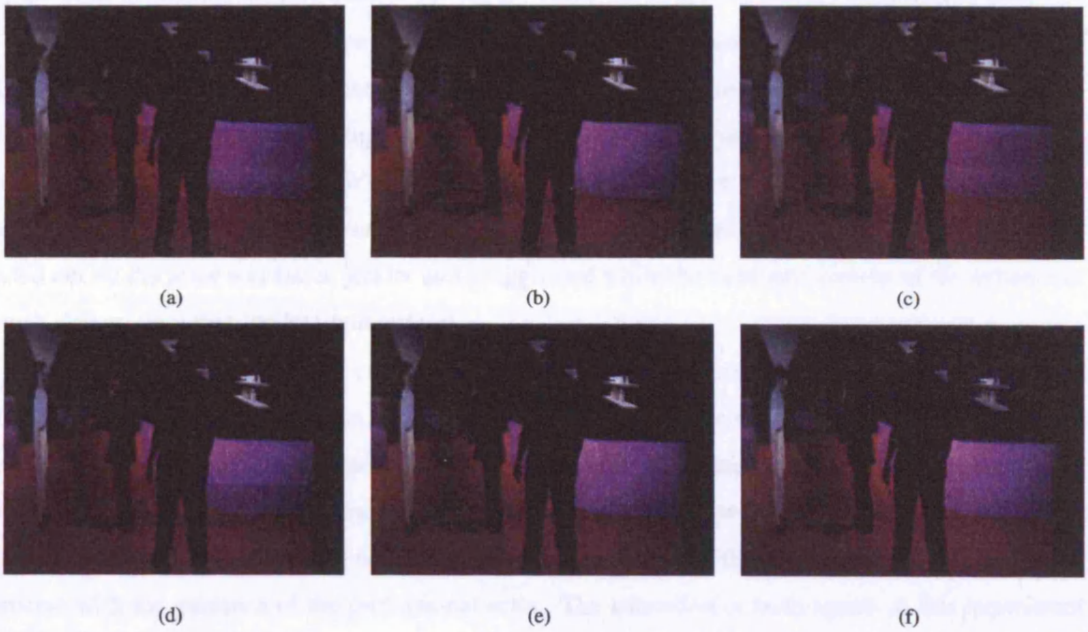


Figure 6.9: The turning process in the real condition

At the start of each of the three sessions, the agents (or actor) acknowledge the participant by turning their heads and looking at the participant. After the acknowledgement, the agents start their conversation. During the conversation, at irregular but sufficiently spaced out times, the agents were scripted to turn their heads and look at the participant. This happened 4 times during each session at about 35 seconds, 65 seconds, 91.2 seconds and 120.2 seconds after the start of the session. Therefore, the agents (or actor) look at the participant a total of 5 times. The manner in which this happened was carefully choreographed so that each time the agents (or actor) looked at the participant, the stimuli perceived by the participant was the same (Figures 6.8 and 6.9). The passive agent which played the role of the listener was scripted to initiate the process. First, the passive agent turned its head towards the participant, after which the active agent turned its head towards the participant. After one section, the passive agent turned its head back towards the active agent followed shortly by the active agent. This process seemed the most natural. In the real conditions, the actor took his cue to turn, towards the participant, from the passive agent and maintained eye contact with the participant for a duration that allowed for the passive agent to be the first to turn its attention back to the conversation.

6.3.6.2 Designing behaviour based on the actor

A process similar to rotoscoping was used to extract key postures and gestures identified by the actor. The actor's improvised performance of the script/context was captured on video tape and imported into 3D Studio Max (Autodesk, 2006). Through multiple meetings with the actor, a set of key postures, gestures and actions were identified for each affect (Angry and Sad). The imported movie file was mapped onto a rectangular plane and placed behind a humanoid skeleton model in 3D Studio Max. Using the actor's performance on tape as a guide, keyframes were created to represent the postures and gestures identified by the actor in addition to other transition postures. These were exported into BVH animation files.

In addition to specific postures such as the bowed head to represent sadness, the importance of the different qualities of body movement such as *speed*, *jerkiness* and *breath* of gestures and actions were discussed with the actor. In keeping with findings presented by Montepare et al. (1999) and Paterson et al. (2001), analysis of the actor's performance revealed that there was a distinct difference in the quality of body movements between an angry performance and a sad performance. An angry motion acted out by the actor was faster, jerkier and exaggerated while the saddened version of the action was much slower, smoother and less pronounced.

A general PIAVCA filter was created to allow for the modification of BVH files in order to change different qualities of the animation. The filter allowed an animation file to be modified spatially and temporally for a set period of time after the beginning and before the end of the movement (Section 6.3.4). This filter was used to define appropriate parameters that allowed for the creation of Angry and Sad motions using the same BVH file. The parameters were defined through an iterative trial and error process with the guidance of the professional actor. The animation of both agents in this experiment were choreographed based on the expertise of the actor and his knowledge of human behaviour.

In addition to the 'neutral' postures (32) used in the previous experiment, some generic postures and gestures were designed such as head nods, tilts and cocks ($\approx$20) and body shifts ($\approx$15). These were used to create animation for the passive agent playing the role of a listener. A total of 60 BVH animation files were created using the key postures and gestures. These were used to create the core animation for the active agent. These animation files were blended with generic body shift motion files in order to create variety. A total of 70 shift motion files were created for this purpose. The next section discusses the parametric model used to generate the active agent's affective behaviour and the listening behaviour of the passive agent.

6.3.6.3 Behaviour generation of the agents

In the *moving virtual* condition, the active agent was designed to display affective postures and gestures while in conversation with the passive agent. The set of animation files used to create the bodily cues were the same for the moving conditions in both the Angry and the Sad set. The only difference was in the background posture of the active agent in the Sad set. The neck and head of the active agent was tilted downwards at an angle of 15° and 25° respectively. In order to avoid having the active agent looking at the floor during the whole conversation with the passive agent, the active agent was also designed to look at the passive agent's face at appropriate times during the conversation. The postures and gestures designed for the active agent were varied by blending them into each other or a body shift motion chosen at random from the set of 70 shift motion files. This was done using existing functionality in PIAVCA. The blend factor was chosen out of a uniform distribution of random numbers between 0.25 and 1. This ensured that not more than 25% of the resultant behaviour was contributed by the shift animation.

The PIAVCA motion filter discussed in Section 6.3.4 was then used to define a set of parameters which were used to manipulate the speed and amplitude of the animation file (Table 6.2). These parameters provided the differences in the quality of body movement between the Angry and Sad moving virtual conditions. As mentioned already, these parameters were chosen after numerous iterations of comparing the resultant animation with the performance of the actor. Each version of the animation was rectified with the guidance of the professional actor. In addition the mean rate at which the active agent displayed postures and gestures was also varied depending on the affective state of the active agent. The rate was controlled by using times generated from an exponential distribution around the pre-defined mean rates (Table 6.2). In the resultant animations, the Angry set was faster, more jerkier and more exaggerated especially when compared to the animation of the passive agent. Similarly, the Sad set was slower, smoother and less pronounced.

	Angry	**Sad**
Speed factor†	*15*	*1*
Amplitude factor	*1.25*	*0.5*
Start delimiter	*0.25 (25%)*	*0.2 (20%)*
End delimiter	*0.05 (5%)*	*0.2 (20%)*
Mean time (ms)	*800*	*1000*

Table 6.2: Parameters used to affect the behaviour of the active agent in the moving virtual conditions.

(a) Static virtual condition (b) Moving virtual condition (c) Real condition

Figure 6.10: The active agent in the Angry set of conditions

(a) Static virtual condition (b) Moving virtual condition (c) Real condition

Figure 6.11: The active agent in the Sad set of conditions

The passive agent was animated using the same algorithm used to animate the agents in the previous experiment (Section 5.3.5.3). The passive agent's behaviour animation was independent of the condition. The postures and animations created for the passive agent were set onto a motion queue using PIAVCA's core functions at intervals generated using an exponential distribution about a mean of 3500 milliseconds. In addition to the bodily cues displayed by both agents, PIAVCA's facial animation functionality was used to create gaze behaviour and blinking animation using the simplified gaze model described in Section 5.3.5.3 of Chapter 5. The gaze animations and the animations which allowed the agents to turn towards the participants at set times, for the previous experiment, were used in the agents for all conditions including the static virtual condition.

	Active Agent	**Passive Agent**
Moving condition	*Affective bodily cues*	*Generic bodily cues*
Static condition	*Static*	*Generic bodily cues*
Real condition	*Real person (actor)*	*Generic bodily cues*

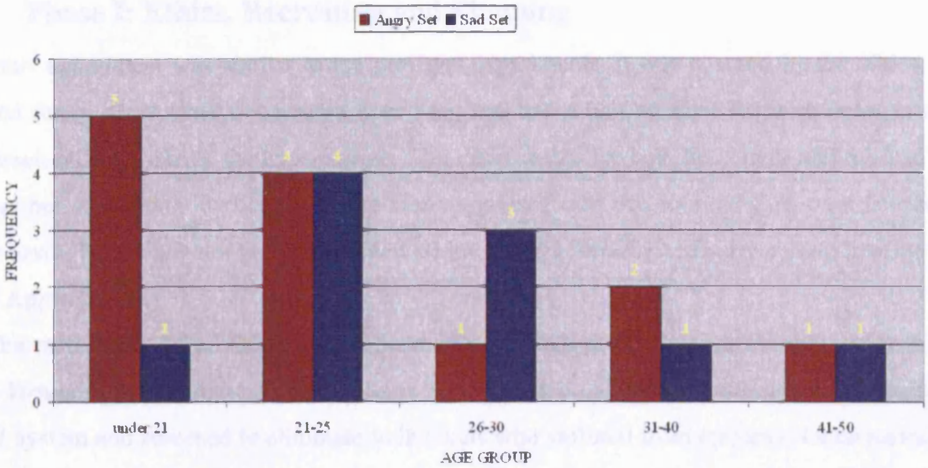
Table 6.3: The behaviour animation of the agents across conditions

6.3.6.4 Audio content of the conversation

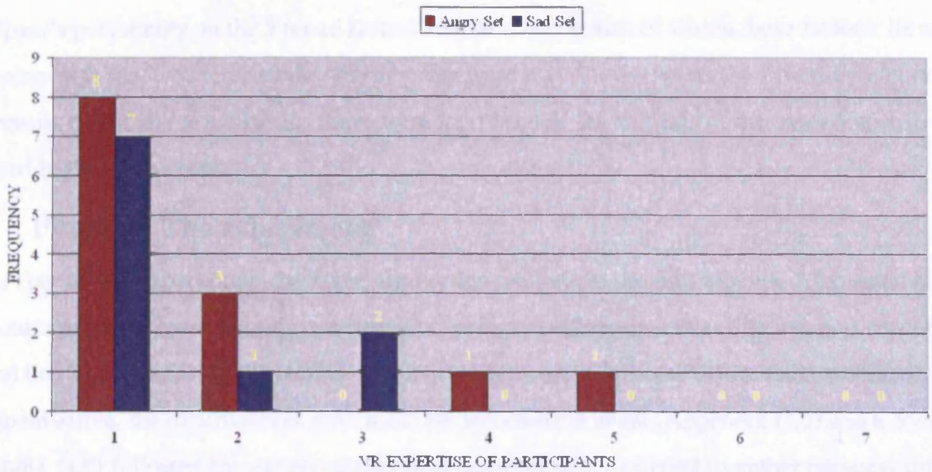
In order to maintain congruent cues, two readings of the script, attached in Appendix G.2, were enacted by the actor in an Angry and a Sad voice. These readings were recorded using the Microsoft Sound Recorder software, converted to the DIVE compatible uLaw format and played back during both the static and moving virtual conditions in order to provide audio content to the conversation between the agents. In the real condition, the actor performed the script live in the appropriate emotional tone while displaying the appropriate bodily cues.

6.4 The participants

Like the experiment discussed in Chapter 5, only male participants were recruited to take part in this experiment. Forty-one male participants were recruited including paid formal pilots. Only twenty-three of these participants undertook the final version of the experiments. Thirteen of these participants were assigned to the Angry set while ten participants were assigned to the Sad set. The data collected from the final 23 participants were analysed.



(a) Age Group



(b) Experience in VR systems (1 - low, 7 - high)

Figure 6.12: Demographics spread of participants

All the participants were under the age of 50 and relatively little experience with VR systems. Eighteen out of the twenty-three participants were students.

		Affect	
		Angry Set	*Sad Set*
Order	*Moving, Static, Real*	7	5
	Static, Moving, Real	6	5

Table 6.4: Overview of population distribution across experiment

6.5 Procedure

There were three phases to the procedures of this experiment as well: preparation, the experiment, and pos-experiment. Experience from the previous experiments were used extensively to efficiently plan this experiment.

6.5.1 Phase I: Ethics, Recruiting and Planning

Since this experiment was similar to the previous experiments, it was covered by the ethics approval obtained for it. Each time slot was an hour long and had a half an hour break in order to allow for unforeseen mishaps during the experiments. The experiments for both the Angry and the Sad set were run together at random. Participants were then assigned to the pre-set time slots on a first come first choice basis. Each time-slot and session was coded using a four-digit identity coding process (Section G.4 of Appendix G).

The recruitment of participants was conducted through poster campaigns only under the project name: *Virtual Bar*. Potential participants were then sent an email giving preliminary information about the VR system and screened to eliminate individuals who suffered from epilepsy. Once partial consent was obtained, the participant was asked to complete an online personality inventory (Johnson, 2006) based on the five-factor model discussed in Section 2.3.5.4. The inventory gave an estimate of the participant's personality on the 5 broad factors and 30 sub-domains of within these factors. Its results are in keeping with the five factor model. Participants were asked to complete the inventory and email back their results before the experiment. These were later used in the analysis of the main response variables collected in the experiment.

6.5.2 Phase II: The experiment

On the day of the experiment, the actor always arrived before the participants. The actor hid behind some curtains in the room housing the ReaCTor until his performance was due in a real condition. This ensured that the participant did not have an opportunity to meet the actor before the experiment. Once the participant arrive, the experimenter gave them the information sheet (Appendix G.5) and a consent form (Appendix G.6) followed by one pre-experiment questionnaire designed to gather personal information (Appendix G.7). Once the apparatus and the participant was ready, the experimenter went through the process of fitting tracking and physiological sensors to the participant as described in Section 3.2.7.

Since this experiment involved exposing the participant to three different conditions, the experiment was conducted in three main stages. Participants were instructed to stand relatively still during the experiment and rigourously still during the baseline periods. Participants were told that the collection of the physiological baseline was important to analysing the results of the experiment. They were also informed that physical motions on their part could add noise to the data. In this way participants were encouraged to stand relatively still in the middle of the ReaCTor (and virtual bar) throughout the baseline periods and sessions.

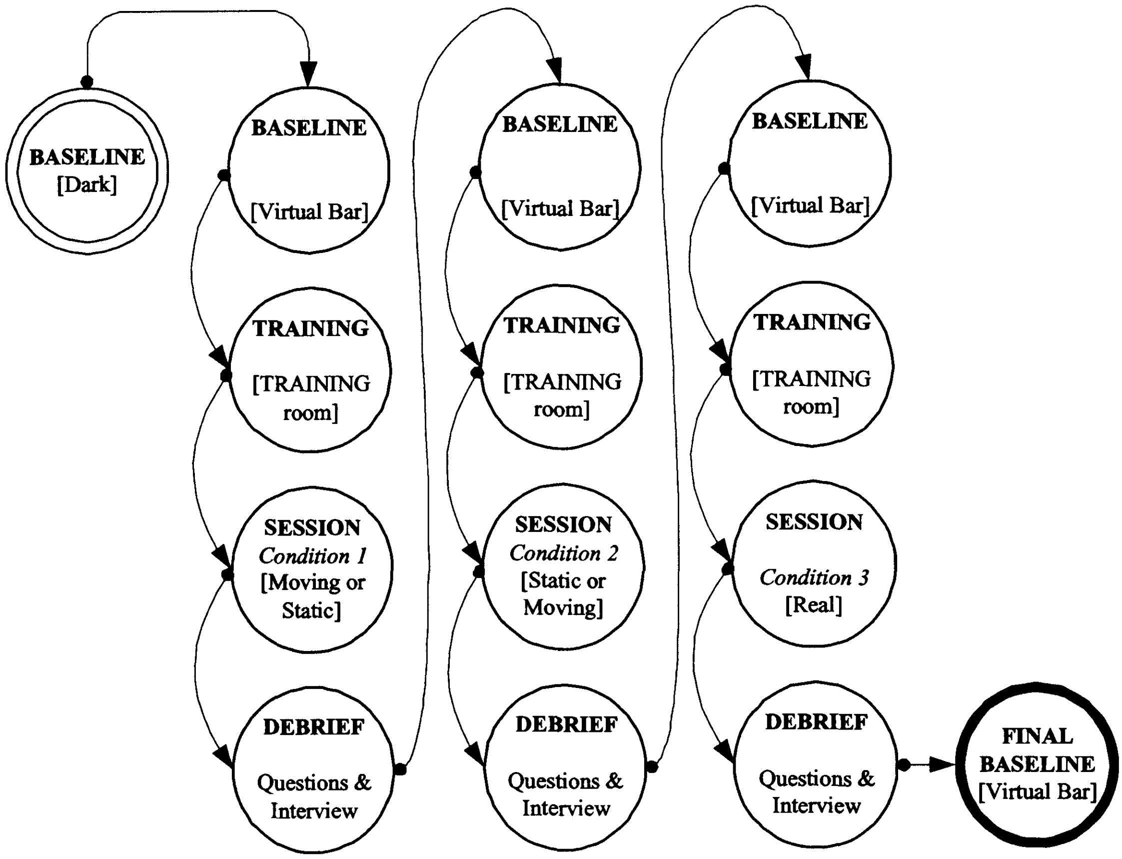


Figure 6.13: Procedure followed in the experiment

Figure 6.13 depicts the procedure followed during the experiment. At the start of the experiments, participants are asked to stand still for in the inert ReaCTor with the lights switched off (darkened room) to record the first of many physiological baselines. Each baseline period lasted for 2 minutes and 30 seconds at least. This period was equivalent to the duration of the sessions. The purpose of maintaining a duration of $2\frac{1}{2}$ minutes for each baseline periods and experiment sessions was in order to allow for the full playback of the chosen soundtrack[1] in the virtual bar. Immediately after the first baseline, the training room was displayed on the walls of the ReaCTor. The participant was then instructed to close their eyes and open them when they heard music. They were instructed that they would find themselves in a virtual bar. They were asked to stand still for a further $2\frac{1}{2}$ minutes in the virtual bar with the music playing in order to take a physiological baseline with any physiological responses to the beats and rhythms of the music. Once the second baseline period was over, the experimenter triggered the training room to be displayed on the ReaCTor. As mentioned in Section 6.3.5, the training room only served the purpose of ensuring that the VR system was working. Participants were give standard instructions.

[1]Each participant listened to the Beatles' version of *"Baby it's you"* seven times during the experiment. The experimenter listened to it at least two hundred and eighty times!

When you hear some music, open your eyes. You will find yourself in the bar with two other people. The person on the right is telling something to the person on the left. Please stand in your spot and observe them. After $2\frac{1}{2}$ minutes, the session will end and we will have a short chat about your experience. This will be repeated three times.

Once the participant closed their eyes, the experimenter triggered the appropriate condition-dependent session. In the virtual conditions, the corresponding active agent and a passive agent were displayed in the virtual bar. The agents were scripted to carry out their pre-determined conversation for about $2\frac{1}{4}$ to $2\frac{1}{2}$ minutes while displaying condition-dependent behavioural cues and occasionally looking a the participant. In the real condition, after the participant closed their eyes, the experimenter signalled to the actor to come out of hiding and stand in position in the ReaCTor. Once the music started, the actor performed the script live in the appropriate emotional tone with the passive agent only. In all the conditions, at the end of the performance/conversation, the virtual bar was rendered invisible thereby darkening the room completely. In the case of the real condition, this gave the actor the opportunity to leave the ReaCTor before the lights were switched on.

(a) Entrance before the start of the session (b) Staging just seconds before the session (c) Leaving (*staged illustration*)

Figure 6.14: The actor's entrance and exit in the virtual bar.

At the end of each session, the participant was administered a section of the questionnaire designed to assess the participant's reported copresence and their judgement of the emotional state of the agents/actor. This was conducted in the ReaCTor. The complete questionnaire is attached in Appendix G.8. Once the questionnaires were completed, the participant was asked to stand still for another baseline period. Only the data collected in the first two baseline periods were used in the analysis, however, the baseline periods helped to ensure that any physiological responses detected during the following sessions were not corrupted. This process was continued until the participant had experienced all three conditions.

6.5.3 Phase III: Post-experiment

All questionnaires and interviews were conducted during the experiment, therefore the only procedures completed in phase III of the experiment were to do with debriefing the participant.

6.6 Measured responses

In the previous experiment, participant responses were recorded using a wide spectrum of measures. Since, this experiment was focused around the participant responses to the active agent in comparison with the real person (actor), two measures were chosen to best suit the research questions: subjective questionnaires and objective physiological measures. Even though the semi-structured interviews used in the previous experiment were useful in obtaining the participant's views on their experience, formal interviews were not conducted in this experiment due to the time constraints. However, the experimenters administered the questionnaires by allowing the participant to explain their responses to the items in the questionnaire. Additionally, participants were asked to describe their observations, of the active character and the agent, after each condition. Since the participants were asked to stand still during the experiment, the analysis of tracking data were not an appropriate measure of participant responses.

The participant's physiological responses were monitored and recorded through all the baseline periods and sessions using the Procomp+ device (Section 3.2.3.3). The questionnaire, attached in Appendix G.8, was completed in three sections - one section per session (condition). Each section of the questionnaire corresponding to the virtual conditions had six items on a 7-point Likert-type scale. The section administered after the real condition contained six items designed to compare participants' subjective assessments of their responses to the active agents between virtual conditions and with the actor. In addition to the copresence questionnaire, participants were asked to judge the underlying emotional state of the agents (and actor) after each session. Participants were asked to choose out of 9 items: *surprised, afraid, angry, happy, disgusted, sad, neutral, don't know* and *other*. The questionnaire was administered in an interactive fashion with the participant in a manner, in the ReaCTor, approaching the style of an informal unstructured interview (Figure G.7 in Appendix G.9).

6.7 Analysis of responses

6.7.1 Findings from Questionnaires

Participant responses collected through the questionnaire were analysed using the same logistic regression method used in all the previous experiments and discussed in Sections 3.4.1.1 and 3.4.1.2.

	Description of variable	**Administered**
Extraversion	*Participant's measure on the FFM: Extraversion*	*Phase I*
Agreeableness	*Participant's measure on the FFM: Agreeableness*	*Phase I*
Conscientiousness	*Participant's measure on the FFM: Conscientiousness*	*Phase I*
Neuroticism	*Participant's measure on the FFM: Neuroticism*	*Phase I*
Openness	*Participant's measure on the FFM: Openness to Experiences*	*Phase I*
Age	*Participant's age*	*Phase II*
Literate	*Participant's level of computer literacy*	*Phase II*
Program	*Participant's expertise in computer programming*	*Phase II*
VR	*Participant's expertise with VR systems*	*Phase II*
Game	*Participant's experience with video game playing*	*Phase II*
Gametime	*Amount of time spent playing video games per week*	*Phase II*

Table 6.5: Explanatory variables used in the experiment

6.7.1.1 Copresence response

	Deviance χ^2	d.f.	Association	$\sim$ P value
Condition • Trial	11.87	1	+	5.70e-004
VR	14.23	1	−	1.62e-004
Neuroticism	46.21	1	+	1.06e-011
Overall	9.7129	20		

Table 6.6: Fitted logistic regression for co-presence response in the Angry virtual conditions

	Deviance χ^2	d.f.	Association	$\sim$ P value
Condition • Trial	8.307	1	+	0.0039
Age	10.39	1	+	0.0013
Extraversion	4.609	1	+	0.0318
Overall	13.260	14		

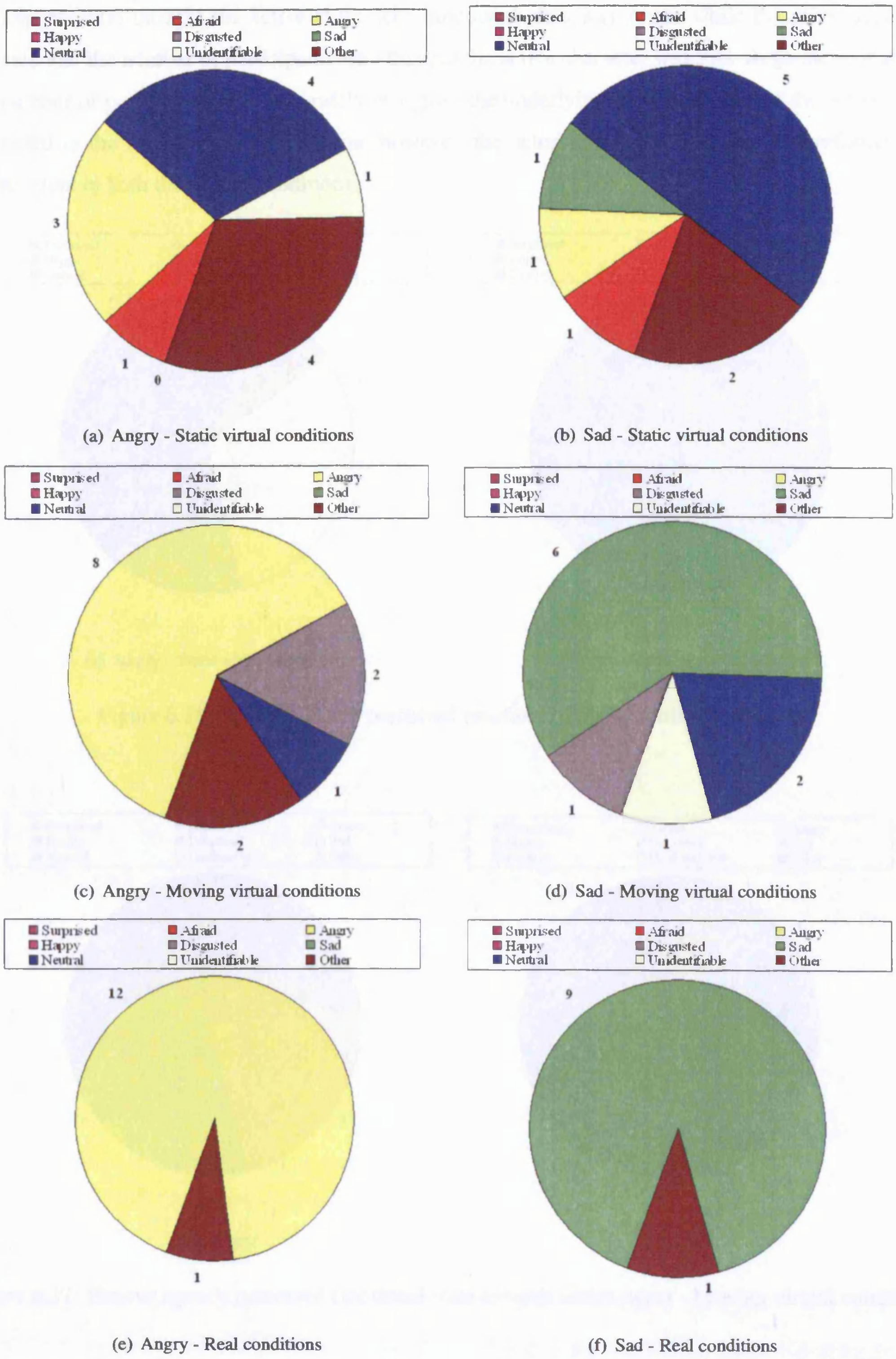
Table 6.7: Fitted logistic regression for co-presence response in the Sad virtual conditions

The same operational co-presence questionnaire used in the previous experiment was modified and used to capture the participant's responses to the agents in both virtual conditions (Appendix G.8). In particular the items were designed to capture if the participant's responses were realistic. The two sets of copresence responses were then analysed using the logistic regression method to uncover any factor-related differences. The two factors were the condition (moving or static virtual condition) and the trial the participant was in (first or second virtual condition - *Trial* number). It is important to note that this was a within-group experiment therefore the variables condition and trial are interconnected. In the fitted model, condition 1 was the moving virtual condition while condition 2 was the static virtual condition. Trial 1 corresponded to the first session the participant experienced and Trial 2 corresponded to the second session.

Independently, the static virtual condition and the second virtual condition in the Angry set was negatively associated reported operational copresence. However, there was a significant interaction effect between the type of virtual condition and the order in which the condition was experienced. When the static virtual condition was experienced after the moving virtual condition, it was positively associated with reported copresence. The same effect was observed in the Sad set. In the Angry set, the participant's expertise with VR was negatively associated with reported copresence while the participant's measure of neuroticism was positively associated with reported copresence. In the Sad set, the age of the participant and the participant's measure of extraversion were both positively associated with reported copresence.

6.7.1.2 Accuracy in recognising underlying emotion

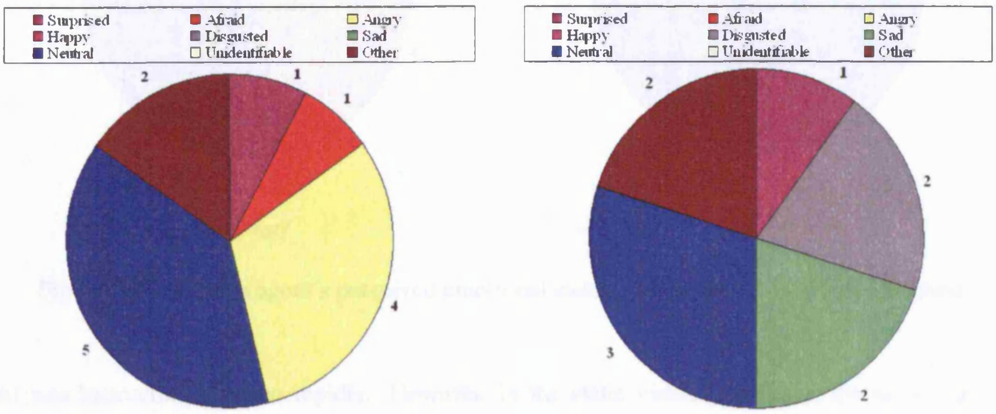
After each session, participants were asked to indicate their judgement of the agents' (and actor's) underlying emotional state. Participants were able to choose from one of seven emotional labels, choose a label of their choice or indicate that they did not know.



(a) Angry - Static virtual conditions (b) Sad - Static virtual conditions (c) Angry - Moving virtual conditions (d) Sad - Moving virtual conditions (e) Angry - Real conditions (f) Sad - Real conditions

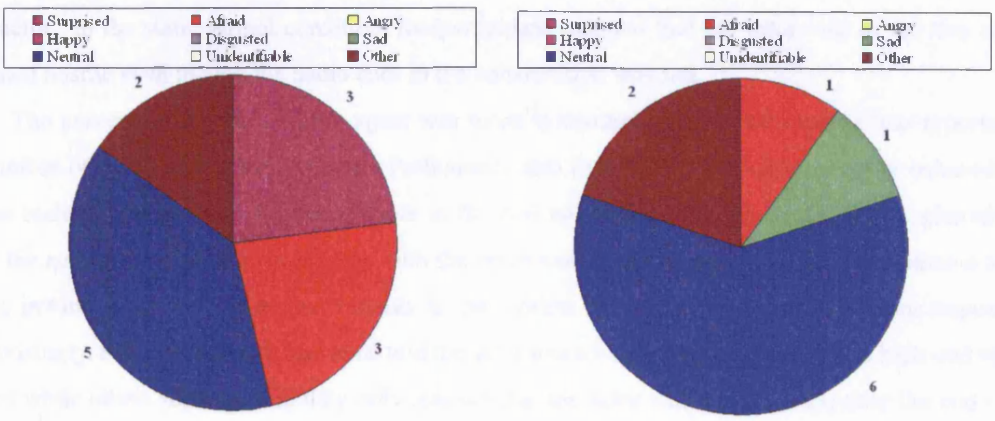
Figure 6.15: Actor's perceived emotional state towards passive agent in both the Angry and Sad sets-Static virtual, Moving virtual and Real conditions

Figure 6.15 show the participant's impression of the active agent's (or actor's) emotional state in both the Angry and Sad conditions. The yellow segments of the pie charts indicate the number of participants who thought the active character's emotional state was Anger while the green segments indicate that the number of participants who thought the active character was Sad. Regardless of affect, the number of participants who accurately recognise the underlying emotional state of the active agent increased in the moving virtual condition, however, the actor in the real condition outperformed the active agent in both the virtual conditions.



(a) Angry - Static virtual conditions (b) Sad - Static virtual conditions

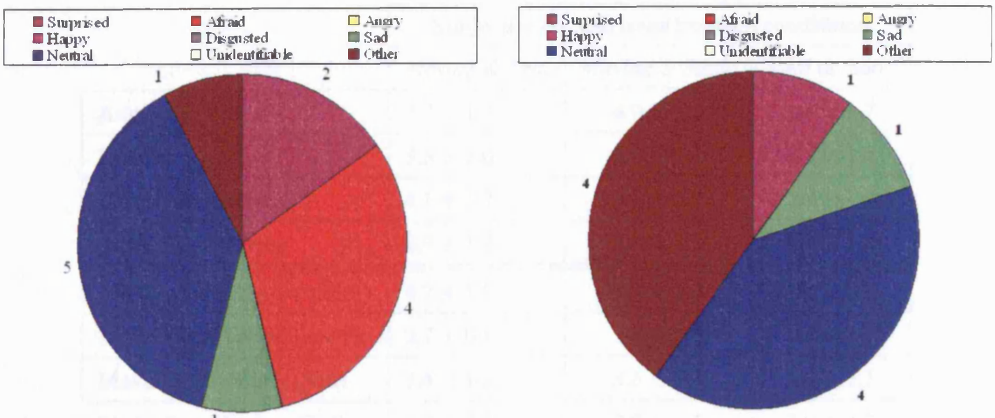
Figure 6.16: Passive agent's perceived emotional state towards active agent



(a) Angry (b) Sad

Figure 6.17: Passive agent's perceived emotional state towards active agent - Moving virtual conditions

Figures 6.16 to 6.18 show the participant's impression of the passive agent's emotional state in both the Angry and Sad conditions. In keeping with the results from the previous experiments, participants judged the emotional state of the passive agent such that it was in keeping with social context they perceived. Participants often reported that the underlying emotional state of the passive agent changed with condition. For instance, one participant (P1107), who experienced the Angry conditions, reported that in the moving virtual condition the passive agent seemed confused and afraid because the active



(a) Angry (b) Sad

Figure 6.18: Passive agent's perceived emotional state towards the actor - Real conditions

agent was instructing him too rapidly. However, in the static virtual condition, the same participant reported that the passive agent seemed more relaxed and was nodding his head while listening to the active agent. Then, in the third and real condition, the same participant reported that the passive agent looked apologetic since the actor was quite aggressive and hostile. In the Sad scenarios, one participant (P2101) thought that the passive agent was simply listening to the active agent in the moving virtual condition, however, in the real condition he reported that the passive agent was understanding towards the actor. In the static virtual condition, the participant reported that the behaviour of the two agents seemed hostile even though the audio cues in the conversation was sad.

The perception that the passive agent was more sympathetic and understanding was reported by a number of other participants as well. Participants also reported that the passive agent behaved in a more realistic manner and responded more in the real conditions with the actor. One explanation is that the actor played his part in keeping with the behavioural cues he perceived from the passive agent. This, in turn, might have given participants the perception that the passive agent was more responsive. Surprisingly, a few participants had to be told the actor was a real person as opposed to a high-end virtual agent while others reported that they only realised that the actor was a real person near the end of the real session.

6.7.1.3 Subjective comparison of the conditions

Participants were asked to compare their responses to the three conditions, through a 7-point questionnaire, at the end of the experiment. Over three-quarters of the participants (78%) reported that the moving virtual condition was the closest to the real condition with respect to the type of realistic responses experienced. Most participants reported that, with respect to the responses felt towards the agents, the static virtual condition was furthest away from the real condition. However, participants did not indicate that their responses in the moving virtual condition and the real condition were exactly the same. Tables 6.8 and 6.9 gives the means and standard deviations of the participant's subjective assessment of their responses to the agents between conditions.

	Subjective comparisons between conditions		
	Moving & Real	*Moving & Static*	*Real & Static*
Angry Set	3.7 ± 1.7	4.0 ± 1.9	2.5 ± 1.7
Sad Set	3.3 ± 2.0	2.9 ± 1.8	2.3 ± 1.4
Moving then Static	4.1 ± 1.7	3.9 ± 2.1	2.9 ± 1.4
Static then Moving	2.9 ± 1.7	3.1 ± 1.7	1.9 ± 1.6
Moving then Static (Angry)	4.7 ± 1.8	4.2 ± 2.4	3.2 ± 1.6
Static then Moving (Angry)	2.7 ± 0.8	3.8 ± 1.5	1.8 ± 1.6
Moving then Static (Sad)	3.4 ± 1.5	3.6 ± 1.8	2.6 ± 1.1
Static then Moving (Sad)	3.2 ± 2.5	2.2 ± 1.6	2.0 ± 1.7

Table 6.8: Means and standard deviations of the participants' assessment of their responses to the agents between conditions.

	Subjective comparisons		
	Moving & Real	*Moving & Static*	*Real & Static*
Angry Set	3.5 ± 1.9	4.8 ± 1.7	3.3 ± 1.9
Sad Set	2.5 ± 1.3	4.8 ± 1.8	2.3 ± 1.4
Moving then Static	3.7 ± 1.7	5.2 ± 1.3	3.4 ± 1.7
Static then Moving	2.4 ± 1.4	4.4 ± 1.9	2.3 ± 1.7
Moving then Static (Angry)	4.0 ± 2.2	5.8 ± 1.0	3.8 ± 2.1
Static then Moving (Angry)	3.0 ± 1.6	3.7 ± 1.5	2.7 ± 1.6
Moving then Static (Sad)	3.4 ± 1.1	4.4 ± 1.3	2.8 ± 0.8
Static then Moving (Sad)	1.6 ± 0.6	5.2 ± 2.2	1.8 ± 1.8

Table 6.9: Means and standard deviations of the participants' assessment of their responses to the agents between conditions specifically when the agents (and actor) turned to look at the participant.

6.7.2 Physiological data: Galvanic Skin Response

The only physiological data analysed was the number of SCRs in the participants' GSR recording and the changes in GSR around the times at which the agents (and actor) looked at the participant.

6.7.2.1 Analysis of SCRs

The number of SCRs in the participants' GSR data for the first two virtual conditions was used as a response variable and analysed in the manner discussed in Sections 3.4.2.1 and 5.7.2.1. The rate of SCRs during the first baseline recorded in the virtual bar was fitted into the final model in order to eliminate any differences in participant physiology.

In keeping with the results obtained in the analysis of the reported copresence responses, separately, the static virtual condition and the second virtual condition in the Angry set was negatively associated with the number of SCRs experienced by the participant. There was also a strong interaction effect between the type of condition and the Trial number. When the static virtual condition was experienced after the moving virtual condition, it was positively associated with the number of SCRs. All the explana-

	Deviance χ^2	d.f.	Association	$\sim$ P value
Condition • Trial	18.95	1	+	1.34e-005
Neuroticism	12.38	1	−	4.34e-004
Openness	18.10	1	+	2.10e-005
Agreeableness	7.058	1	−	0.0079
Extraversion	10.76	1	−	0.0010
Conscientiousness	7.049	1	+	0.0079
Overall	11.879	8		

Table 6.10: Fitted linear regression for the number of SCRs in the Angry conditions

tory variables associated with the personality of participants were highly significant and were therefore included in the log-liner regression model (Table 6.10). The participant's measure of neuroticism, agreeableness and extraversion were negatively associated with the number of SCRs while the participant's measure of openness and conscientiousness were positively associated with the number of SCRs.

	Deviance χ^2	d.f.	Association	$\sim$ P value
Condition • Trial	5.191	1	+	0.0227
Neuroticism	13.08	1	−	2.98e-004
Openness	38.55	1	−	5.34e-010
Extraversion	13.67	1	−	2.18e-004
Conscientiousness	6.671	1	+	0.0098
Overall	18.199	9		

Table 6.11: Fitted linear regression for the number of SCRs in the Sad conditions

In the Sad sets, the negative association between the static condition or the second condition and the number of SCRs was not significant. However, like the Angry set, there was a strong interaction effect between the factors. With the exception of the participant's measure of agreeableness, all other personality measures were significantly associated with the number of SCRs experiences (Table 6.11). The participants' measure of conscientiousness was positively associated with the number of SCRs while participants' measure of neuroticism, openness and extraversion were negatively associated with the number of SCRs.

6.7.2.2 Event-related analysis

The moments at which the agents (or the actor) turns to look towards the participant during the experiment are referred to as '*events*'. In order to assess if the act of turning to look at the participants elicited any responses, event-related analyses were carried out on the participants' GSR recordings. As discussed in Section 3.4.2.1, a window of GSR data ($\pm$ 10 seconds) corresponding to all the events experienced by the participant is extracted and averaged to detect if there are any characteristic patterns or changes in the GSR data. Figure 6.19 shows the GSR data averaged around all the events experienced by all the participants. Time 0 is the moment at which the event occurs. The green curve shows the physiological response experienced by the participants when the passive agent and actor turns to look at them. This

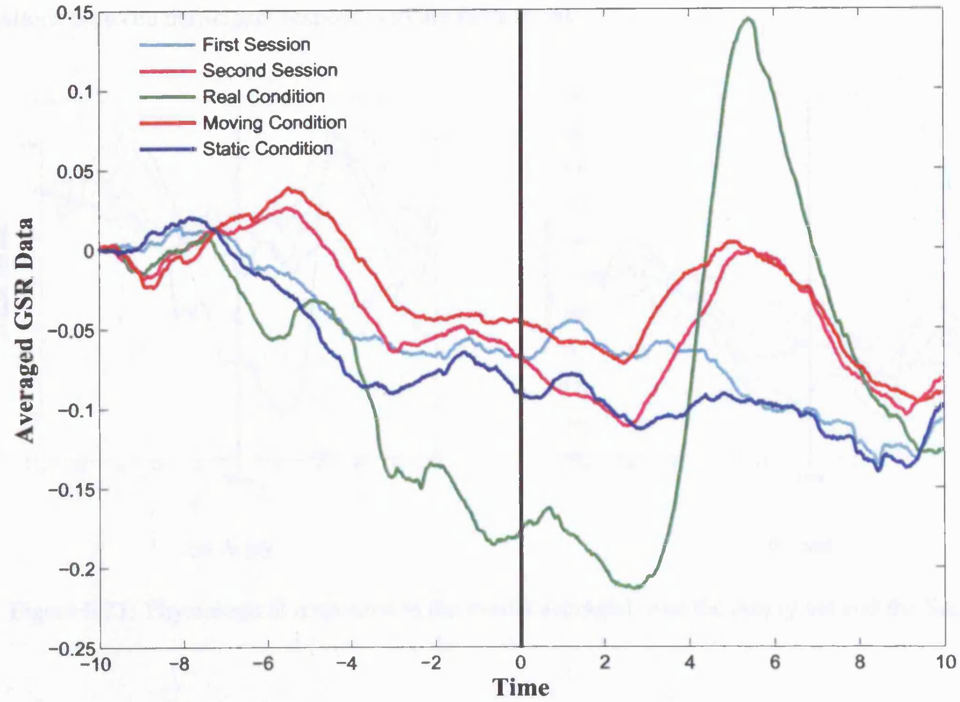


Figure 6.19: Response to the events averaged over all the participants.

peak in the GSR data occurs in the moving virtual conditions and the second virtual conditions as well but not to the same extent.

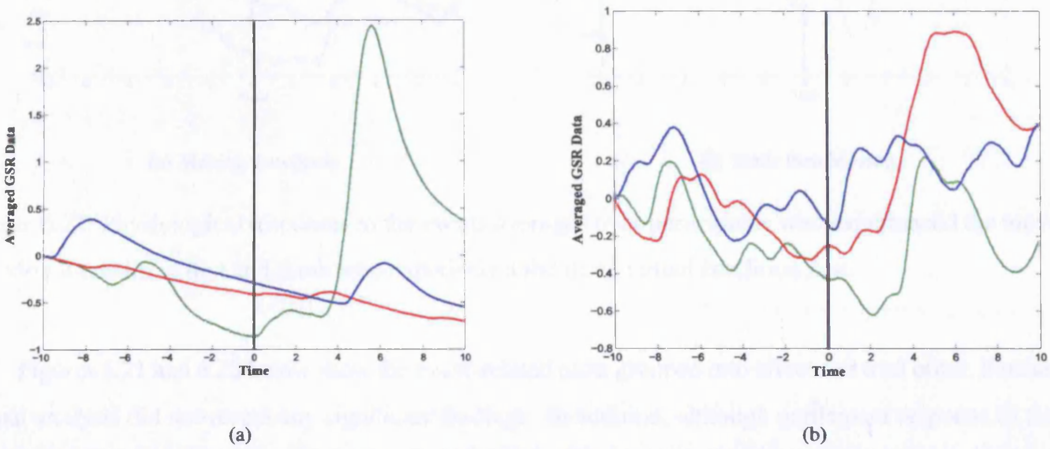
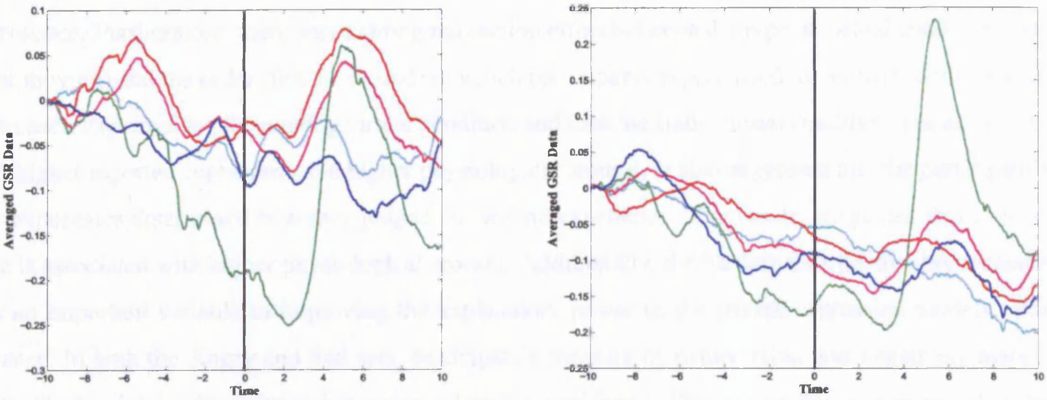


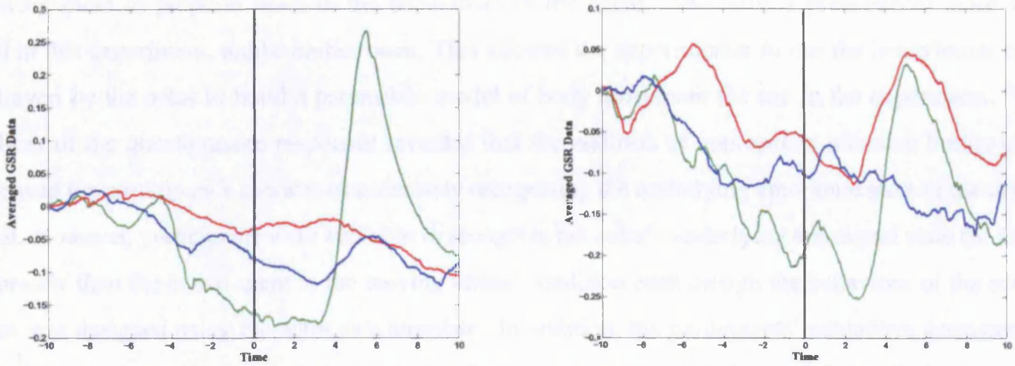
Figure 6.20: Physiological responses to the events in the virtual bar for two different participants.

Figure 6.20 shows responses from two participants. The green curves represent the participant's response to seeing the passive agent and the actor turn to look at them in the real condition. The blue curves represent the participant's response to seeing the active and passive agents turning to look at the participant in the static virtual conditions. The red curves represent the participant's response to seeing the agents turning to look at the participant in the moving virtual conditions. The figures show the variations between participant responses to the same event.



(a) Angry (b) Sad

Figure 6.21: Physiological responses to the events averaged over the Angry set and the Sad set.



(a) Moving then Static (b) Static then Moving

Figure 6.22: Physiological responses to the events averaged over participants who experienced the moving virtual condition first and those who experienced the static virtual condition first.

Figures 6.21 and 6.22 below show the event-related plots grouped into affect and trial order. Further visual analysis did not reveal any significant findings. In addition, although participant response to the event in the real conditions were more pronounced, this effect was magnified by two participants (Figure 6.20a) who were especially non-responsive to the virtual conditions but were incredibly responsive to the real condition.

6.8 Discussion

This experiment was designed to evaluate the impact of affective bodily cues on participant responses. Like the previous experiments (Chapter 5), the affective states investigated were Anger and Sadness. Two types of participant responses were gathered during the experiment: questionnaires and physiological arousal.

In both the Angry and Sad conditions, a static virtual condition was associated with reduced reported copresence. Furthermore, there was a strong interaction effect between the type of virtual condition (static or moving) and the order (first or second) in which participants experienced the virtual conditions. In both cases, experiencing the moving virtual condition and then the static virtual condition was associated with higher reported copresence and higher physiological arousal. It also suggested that the participant's first experience determined how they judged the second experience. The results suggested that copresence is associated with higher physiological arousal. Additionally, the participant's personality measure was an important variable in improving the explanatory power of the overall regression models. For instance, in both the Angry and Sad sets, participant's measure of extraversion was negatively associated with the physiological arousal experienced by the participant. This association was mirrored in the participants' measure of neuroticism as well.

In the previous experiment, one of the main obstructions faced in interpreting the data was the lack of a baseline. This was solved in this experiment in two ways. Firstly, the experiment employed a within-group design. This allowed participants to compare and contrast between conditions, thereby, allowing them to pinpoint flaws in the behaviours of the agent. Secondly, a professional actor was used in this experiment, unlike earlier ones. This allowed the experimenter to use the behavioural cues portrayed by the actor to build a parametric model of body movement for use in the experiment. The analysis of the questionnaire responses revealed that the addition of appropriate affective bodily cues improved the participant's chances of accurately recognising the underlying emotional state of the active agent. However, participants were still able to recognise the actor's underlying emotional state far more accurately than the active agent in the moving virtual condition even though the behaviour of the active agent was designed using the actor as a template. In addition, the participants' subjective assessments of their responses to the agents in between conditions suggest that participants felt that their responses to the real conditions were more similar to the moving virtual conditions in comparison to the static virtual conditions. This suggests that there are some undefined parameters that need to be included to the parametric model of affective body movement presented in this Chapter.

Participants also mentioned hearing the actor clapping his hands together when gesturing empathically in the Angry condition. This helped the participants to discern the emotional state of the actor. Furthermore, the actor's behavioural cues were in sync with the audio cues unlike the active agent in the moving virtual condition. This was reflected in some of the comments made by participants. It is quite possible that a more congruent and synchronised parametric model of multiple behavioural cues would be sufficient to overcome these issues.

The addition of a real condition in the within-group design of the experiment also allowed for the recording of a physiological template of how the participant would react to a real person. This was used to determine if the participant reacted to the active agent in the virtual conditions as they would in the physical world. Event-related analysis was carried out on the physiological responses of the participants.

The analysis focused on changes in physiological recordings when a significant event occurred. In this case, the event referred to the moments when the agents (and the actor) looked at the participants. In previous experiments reported in Chapters 4 and 5, eye contact and gaze behaviour played an important role in making the agents seem aware of the participant. In keeping with the results from the previous studies, many participants mentioned that the event was significantly arousing especially the first time round. The event-related analysis revealed that the agents in the moving virtual condition elicited a greater response than the agents in the static virtual condition. However, the response was not as profound as the one elicited in the real conditions.

In addition to the event-related analysis, participants were asked to compare and contrast their responses in-between the moving virtual, static virtual and real conditions. Although participant responses in the moving virtual condition were judged as being closer to the real condition in comparison to the static virtual condition; the participant reported that their event-related responses in the moving virtual condition was closer to the static virtual condition than the real condition. This highlights the importance of eye contact in the physical world. However, overall the highly significant findings suggest that, in addition to accurately portraying the emotional state of the active agent, the parametric model of affective bodily cues presented in Section 6.3.6.3 had an impact on participant responses. This suggests that it is possible to accurately portray the emotional state of an agent in IVE through postures and body movement at least with respect to Anger and Sadness.

Finally, in keeping with the results of the previous experiment discussed in Chapter 5, participants have a propensity to interpret the behaviour of an agent using cues perceived from other agents or the virtual environment in a holistic fashion. A number of participants reported that the behaviour of the passive agent changed over the three conditions they experienced. A few went further to state that the passive agent seemed more "*real*" and understanding in the presence of the actor. This phenomena is interesting since it occurred in both the experiments in which a dormant and non-evolving agent was placed simply to create a context which facilities the portrayal of strong emotions. This suggests that 'neutral' cues can indeed be used to portray different scenarios in a VE simply by changing one aspect of the scene.

6.9 Summary

This chapter presented an experiment designed to explore two research challenges. The first was to use the findings from the experiment on posture, discussed in Chapter 5, to develop a parametric model of affective kinesic behaviour (body movement) for use in the experiment. The second was to determine if participants reacted to affective agents in IVE in the same way they would have, had the agent been a real person. A novel approach was adapted by hiring an actor to play the part of an agent in the virtual bar for one of three conditions in a within-group experiment (Section 6.3).

Participants were asked to observe three different versions of two characters having a conversation in an affective state. Two affective states were chosen for investigation: Anger and Sadness. Previous studies by Coulson (2004), Montepare et al. (1999) and Paterson et al. (2001) suggested that the qualities of body movement associated with Anger and Sadness were sufficiently different that people in the

Virtual Human	Agents
Factors	Type of behaviour (Body movement) - Single subject design with three levels
Participants	23 males (Angry - 13 and Sad - 10)
Environment	Virtual Bar
Apparatus	ReaCTor Physiological devices A professional actor
Software	DIVE VRPN (C++ Plugin) PIAVCA (C++ Plugin) PIAVCA Motion Filter (C++ extension)
Data Collected	Questionnaires Physiological measures
Data Analysed	Questionnaires Physiological measures

Table 6.12: Summary of the experiment on body movement

physical world do not confuse one emotional state for the other. Two of the three conditions (versions) participants were asked to observe were virtual conditions: static and moving. The third condition was a real condition in which participants observed a real person (an actor) having an affective conversation with a passive virtual agent. In the two virtual conditions, an affective active virtual agent was seen having the same affective conversation with a passive virtual agent. In one of the two virtual conditions (static), the affective virtual agent was static. In the second of the virtual conditions (moving), the affective active agent was animated using a parametric model of bodily cues which manipulated the *speed, jerkiness* and *breath* of animations to portray, depending on affect, either an Angry virtual agent or a Sad one.

Sections 6.4 to 6.5 discussed the population distribution and procedures used in the experiment. Section 6.6 focused on the methods used to measure two types of participant responses: reported copresence and physiological arousal. Section 6.7 dealt with the analysis of the collected data. The findings suggested that the parametric model increased the sense of reported copresence and physiological arousal experienced by participants in comparison to the static virtual condition. However, a strong interaction effect between the conditions and the order, in which the participant experienced the conditions, meant that when participants were exposed to the static virtual condition after the moving virtual condition, they experienced higher reported copresence and higher physiological arousal. Further experiments are needed to investigate this effect. A discussion of the findings was presented in Sections 6.8. Section H.1 in Appendix H gives a list of the main findings in this experiment.

The next chapter will consider the research questions posed in the first chapter of this thesis and discuss how the experiments reported in this and the previous chapters (Chapters 4 & 5) contributed towards answering them.

Chapter 7

Conclusion

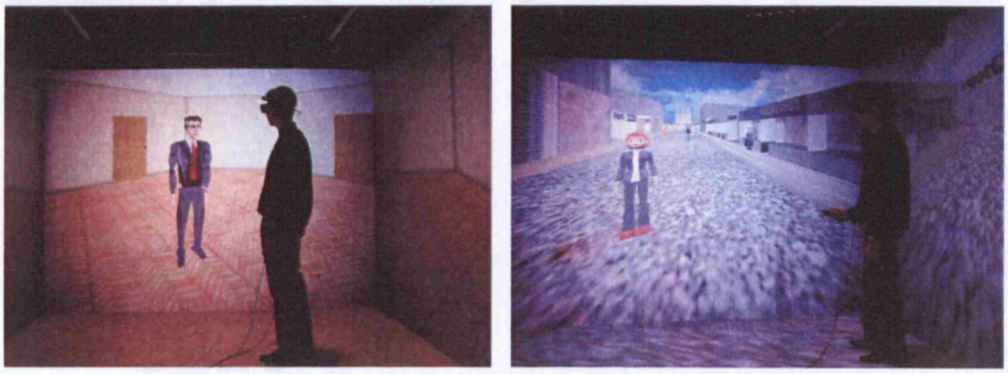
Virtual humans are an essential part of many applications ranging from virtual therapy to games. They are especially suited to applications that require communication since they allow for potentially natural participant-virtual human interaction. However, the visual appearance of a virtual human automatically invokes expectations in participants regarding how the virtual human should behave, to such an extent that any flaws in the virtual human's behaviour can reduce its impact on participant responses. It has been argued that participants will not respond appropriately to a virtual human, if the behavioural capability of the virtual human is not up to the participant's expectations (Slater and Steed, 2002). On the other hand, empirical work suggests that participants do respond in a realistic manner to virtual humans with even minimal behavioural cues (Freeman et al., 2003). Furthermore, parametric models of some behaviours have been successful in eliciting realistic participant responses. For instance, models of gaze behaviour have proven effective in improving participant's perceived quality of communication (Deng et al., 2005; Garau et al., 2001). This suggests that building expression into virtual humans through the use of parametric behaviour models can potentially enhance their effectiveness.

Creating convincing expression in virtual humans, for use in shared immersive virtual environments, presents significant challenges. One of these challenges involve defining an underlying psychological state in virtual humans. As discussed in Chapter 2, research into building affective virtual humans can be divided into two distinct but inter-related themes: the *generation* of the virtual human's internal states and the *expression* of these states through behavioural cues. This thesis has focused on the later of these two themes. The research presented in this thesis investigated the impact of parametric models of virtual human behaviours on different participant responses in an IVE. The main premise behind this thesis is that key attributes associated with a psychological state can be used to invoke realistic responses in participants where responses vary from physiological responses to social and cognitive responses (Sanchez-Vives and Slater, 2005). However, one of the technical challenges faced in this thesis, was the lack of parametric models that could be used to define key behavioural attributes (minimal cues) which are associated with a particular psychological state.

A series of four experiments was conducted using parametric models of gaze and kinesics to determine ways of modelling expression in a full-body virtual human. In order to test the effectiveness of using parametric models of behaviours in an IVE, the first two preliminary experiments were designed to investigate the importance of behaviour fidelity relative to the visual appearance of virtual humans (Chapter 4). The remaining two main experiments were designed to investigate the importance of kinesics in the communication of affect. The literature review (Section 2.4 of Chapter2) revealed that the role of bodily cues in the communication of an affective state was less studied in comparison to facial expression (de Gelder, 2006). Therefore, a parametric model of affective bodily cues was designed for two specific emotional states: Anger and Sadness. The parametric model was used to create two affective virtual humans. Apart from building viable parametric behaviour models, a secondary challenge in the thesis concerned the methods used to evaluate the effectiveness of the behaviour model. A combination of subjective and objective measures used to explore different levels of participants' responses to the affective virtual humans (Chapter 3).

Sections 7.1 to 7.3 summarises the four experiments presented in this chapter in addition to the main findings. Section 7.4 presents the contributions made through this thesis. Finally, Section 7.5 presents future work to extend this thesis.

7.1 Preliminary experiments on perceived behavioural fidelity



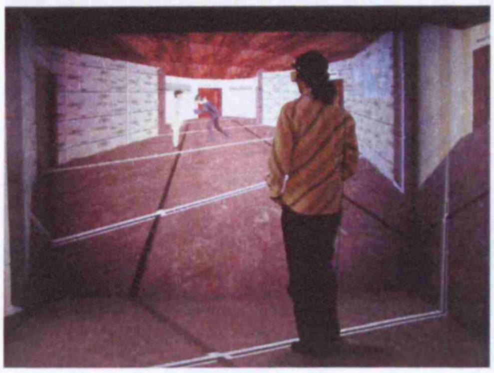
(a) On aspects of avatar fidelity (b) On visual realism

Figure 7.1: Revisiting the preliminary experiments

The first two preliminary experiments, designed to investigate the impact of behavioural fidelity in virtual humans, were reported in Chapter 4. The experiments were designed to assess the extent to which virtual humans controlled by parametric models of behaviours elicit appropriate responses. Each experiment tackled aspects of the research question in a different manner. The first preliminary experiment was designed to explore the impact of two related aspects of avatar fidelity (visual appearance and behavioural fidelity with respect to single behaviour eye gaze) on the participant's perceived quality of communication. The findings from the experiments have already been reported by Garau (2003), however, the parametric model of gaze behaviour used in the experiment was discussed in this thesis (Section 4.1.5 of Chapter 4). The second preliminary experiment explored the different responses to virtual humans, of very low behavioural complexity and limited expression, as reported by participants.

Findings from the first preliminary experiment on avatar fidelity revealed that avatars with higher visual realism needed to command higher behavioural fidelity in order to enhance the perceived quality of communication in IVEs even when the behaviour concerned was as subtle as gaze behaviour. The main contribution from this set of experiments was the parametric gaze behaviour model for dyadic situations. The gaze model was based on a number of parameters which defined the speed of the eye, the direction the eye moved in and the time it spent in a particular position based on the conversation in a dyad. Participant responses to the texture-mapped avatars of their conversational partner were enhanced when the gaze behaviour of the virtual human was controlled by the parametric model. Findings from the second preliminary experiment revealed that independent of the visual realism of the agent, the most significant factor positively associated with reported presence was the participant's perceived sense of realism regarding the behaviour of the agent. The preliminary experiments exemplified the potential advantages of using a parametric behaviour model to enhance participant responses to virtual humans in an inexpensive manner and served to re-enforce the importance of behavioural fidelity in virtual humans.

7.2 Experiment on posture



(a) Approach period

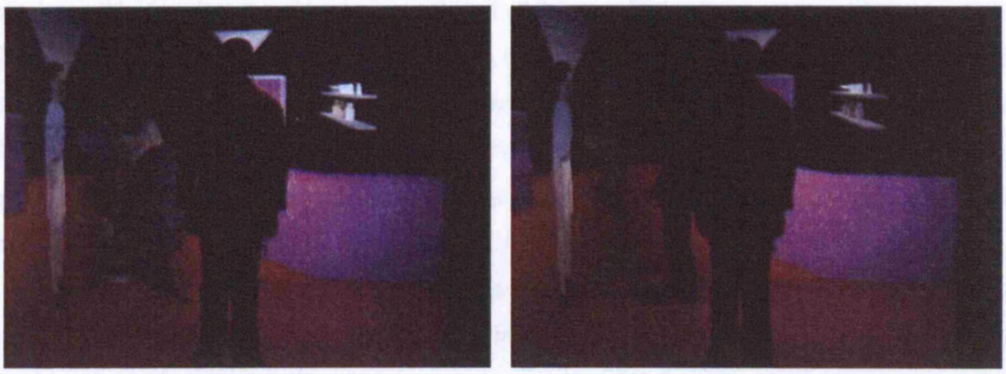
(b) Participant-agent interaction period

Figure 7.2: Revisiting the experiments on posture in the virtual maze

The goals of the experiment presented in Chapter 5 was to evaluate a range of participant responses to sequences of postures exhibited by agents - in particular the focus of interest was in investigating if participants responded appropriately to affective postural cues displayed by a virtual human. The main technical challenge encountered involved designing plausible affective bodily expressions for virtual humans. The first revolved around the lack of systematic studies describing a set of behaviours for an emotional state as highlighted in Chapter 2. In both the Affective (Angry and Sad) conditions, the neutral postural cues were positively associated with reported copresence. In the Angry conditions, the neutral postures were associated with higher physiological arousal. However, in the Sad conditions neutral postures were associated negatively with physiological arousal. Even in the Angry conditions, where participants recognised the active agent's emotional state as Anger, the participants responded with higher levels of copresence and physiological arousal to the 'Neutral' postures. This suggested that although postures play a significant role in the communication of affect, the affective posture model used in the experiment was incorrect.

Similar to the findings in the second preliminary experiment, participants interpreted the social context presented in the virtual maze holistically and attributed the agents, especially the neutral passive agent, with a wide variety of emotional states, attitude and status. Qualitative analysis of the participant interviews revealed that, despite having neutral cues, participants still attributed a wide variety of perceived emotional states or attitudes to the agents. This is in keeping with findings reported in Freeman et al. (2003), where participants attributed different internal states and sentience to simple behaviours displayed by agents. This experiment revealed the participants' surprising propensity to a) interpret any sort of perceived agent expression as meaningful and b) respond socially to the agents. This suggested that a virtual human controlled by an incorrect parametric model of posture prevents participants from experiencing realistic responses.

7.3 Experiment on kinesics



(a) A moving virtual condition (b) A real condition

Figure 7.3: Revisiting the experiments on kinesics in the virtual bar

The experiment, presented in Chapter 6, was designed to extend the experiment on posture to include the role of body movement in the portraying two affective states: Anger and Sadness. An actor was used as part of building a virtual scenario. This solved the main technical challenge faced in building a parametric model of affective postures and body movement. In addition to the existing research presented by Montepare et al. (1999) and Paterson et al. (2001), numerous consultations with the actor was made in building a parametric model of affective body movement.

Three main parameters (*breath*, *jerkiness* and *speed*) were used to define an Affective set of behavioural cues using the same set of animation files. Like the findings presented by Montepare et al. (1999) and Paterson et al. (2001), the Angry motions were faster, more jerkier and more exaggerated in comparison to the Sad motions. A lot of highly significant findings, uncovered in the experiments, were shared by both the Angry and the Sad conditions. Independently, the static virtual condition and the second virtual condition were both negatively associated with reported copresence, however, an interaction effect was uncovered. Participants responded in a favourable manner if they experienced the moving virtual condition followed by the static condition. This interaction effect was carried onto the participants' physiological arousal.

In addition, the participants' ability to accurately recognise the underlying emotional state of the affective agent was enhanced in the moving virtual condition. However, the actor in the real condition clearly outperformed the affective agent in the moving virtual condition especially when the responses were related to direct interaction with the participants. In keeping with findings by Paterson et al. (2001) and Montepare et al. (1999), the quality of body movement played a highly significant role in the communication of affect. The parametric model was successful in eliciting appropriate participant responses. However, there are still some more missing parameters which can better describe the quality of affective body movement with regards to Anger and Sadness. The analysis of participant responses in this manner proved to be a very effective manner in which to design behaviour models for virtual humans.

7.4 Contributions

While research into the visual representation of virtual humans has come a long way, the creation of expressive and interactive virtual humans capable of eliciting realistic responses in participants of virtual environments remains a distant goal. This is often the result of a lack of variety in behaviours, a lack of emotional range, a lack of evolution in its interaction with the participant or a lack of perceived intelligence/understanding. However, creating an affective virtual human presents significant challenges. One of the main hurdles encountered in creating expression virtual humans is the overwhelming amount of information on factors that affect expression and responses to expressions.

Chapter 2 discussed a definitive overview of the models and methods used in the creation of expression and a perceived psychological state in virtual humans. The review focuses on two perspectives of expression: the *technical modelling of interactive expression* and *the theoretical models of psychology* that provide a framework for understanding expression and responses. The chapter gave an overview of a number of areas including: the functions of nonverbal communication, theoretical background and existing models of emotions and personality, the different modalities used to express an internal psychological state, the possible responses that can be elicited by an individual in the physical world and a short description of the concept of presence. The review suggest that all though there are certain characteristic attributes for affective states, mapping those behavioural attributes/cues to a specific affective state is not straightforward.

In comparison to data-driven approaches of building expression in to virtual humans, parametric models of behavioural cues are compact and flexible. Two parametric behaviour models were developed and tested successfully in the experiments discussed in Chapters 4 to 6. The first behaviour model concerned gaze behaviour in a dyadic social context and was used extensively in the first preliminary experiment (Section 4.1.5.1). The parametric gaze model was modelled based on a simple parametric model used by Garau et al. (2001) and a data-driven model reported by Lee et al. (2002). In an preliminary evaluative experiment, participant responses to avatars controlled by the gaze model were enhanced if the avatar was of higher visual fidelity. This suggested that behavioural fidelity plays an important role in eliciting participant responses. The second parametric model focused on affective kinesics. The model was built using inferences from previous works, such as those presented by Coulson (2004) and Paterson et al. (2001), and a pseudo-data driven process. The data-driven process involved observing the behavioural cues used by a professional actor - Gregory (2006). A parametric model was designed to reflect an Angry and Sad emotion state using body movement and a few key postural attributes (Section 6.3.6.3). The model of body movement was successfully used in the experiment on kinesics in conjunction with a simplified version of the gaze model.

This thesis investigated the impact of behavioural fidelity on participant responses with regards to virtual humans in an IVE using the two parametric models mentioned above. The parametric model of kinesics was used in an experiment to investigate if a) participants accurately recognised the underlying affect of a virtual human as portrayed using the model, and b) if participants responded appropriately to the affective virtual human. The parametric model was controlled through three parameters which altered the spatial and temporal properties of the same set of animations to express Anger or Sadness in a virtual human. Results from both experiments revealed that postures and kinesics do play an important role in the communication of affect in IVEs. Participants were able to accurately recognise the underlying emotional state of the virtual human. In addition, participants responded appropriately to the affective virtual human. This finding was supported through both objective and subjective measures. This suggests that affective kinesic cues can be modelled successfully through the use of parametric models.

One of the basic research questions investigated within this thesis focused on whether an expressive virtual human can elicit appropriate responses in participants within an IVE. The first preliminary experiment on avatar fidelity provided findings to support the view that even minimally expressive virtual humans can indeed affect participant responses in a positive manner. This is in keeping with previous studies in which participants attribute sentience and experience paranoia in response to virtual humans with minimal behaviour Freeman et al. (2003). The second preliminary experiment on visual realism revealed that most significant factor associated with higher reported copresence was the participant's perception of the behavioural fidelity of the virtual humans.

Findings in the experiments on posture and kinesics also reflected that participants respond differently at a physiological level to various affective behavioural cues presented to them. For instance, findings from the experiment on kinesics revealed that participants are more physiological aroused in the presence of a virtual human displaying angry bodily cues in comparison to a static virtual human. A higher physiological arousal in response to angry (threatening) behavioural cues is inline with studies by both Luborsky et al. (1963) and Meehan (2001) on threatening pictures and stressful virtual environment respectively. In addition to physiological and subjective responses, most participants maintained the appropriate interpersonal spacing with the virtual humans in the experiment on posture. This was especially the case during the first participant-virtual human meeting. This is in keeping with arguments made by a number of other researchers, including Reeves and Nass (1996) and Pertaub et al. (2002), that participants treat virtual humans as social entities.

Findings from the experiments suggest that participants interpret a social context portrayed to them in the IVE in a holistic fashion. Participants generally use cues around the virtual human in addition to the virtual human's nonverbal behaviour to interpret the emotional state and persona of a virtual human. For instance, when asked to judge the two virtual humans in the experiment on kinesics, one of which was designed to be angry, participants often judged the emotional state of the other virtual human as afraid or surprised. Even in cases where the virtual human has extremely low expressive abilities like in the second preliminary experiment, participants often attributed attitudes of unfriendliness or grumpiness to the virtual human. This is in keeping with James's (1932) premise that one aspect of an expression colours the perception of the whole expression.

7.5 Potential Directions for Future Work

Within the scope of this thesis, virtual humans encompass entities that are designed to appear humanoid from cartoonish to visually sophisticated and from agents to avatars. As discussed in Section 1.3 of Chapter 1, the categorisation of a virtual human along the agent-avatar continuum depends on how the behaviour of the virtual human is controlled. The contributions of this thesis apply to both agents and avatars. In agents, the behaviour of the virtual human is defined through a variety of pre-scripted means. In avatars, the participant has a measure of control over the behaviour of the virtual human. The behavioural modelling of the avatar is not relevant if the participant has absolute control over the avatar's behaviour. However, in cases where a participant has limited tracker control over their avatar, the avatar's behaviour can be inferred through the parametric modelling of behavioural cues controlled through high-level psychological states. In these instances, the challenges faced in representing the autonomous agent's or participant's psychological state are the same. There are other issues involved in building behavioural models for avatars. For instance, participants of virtual environments might not want to represent their true psychological states through their avatars in the same way individuals in the real world might not want to express a certain emotional state in public. These issues are beyond the scope of this thesis and can only be truly explored using an integrated model that deals with both the generation and representation of psychological states in virtual humans.

The contributions listed in this thesis have implications for the study of virtual humans in the fields of both computer science and social psychology. In the fields of computer science, the advantages of modelling virtual human behaviour have been discussed in Section 2.1 of Chapter 2. From a social psychology perspective, the parametric modelling of behavioural cues in virtual humans allows researchers to study real human behaviours with an unlimited level of control. Furthermore, the use of a variety of participant responses to evaluate the effectiveness of the behaviour model allows researchers to study (and isolate) specific behavioural attributes that are important to portray a pre-defined psychological state. The findings from these types of evaluative studies can then be applied to building virtual humans with high-fidelity behaviours for use in virtual environments. The parametric model of gaze behaviour (and any other facial cue) has implications for the creation of virtual human behaviour in virtual environments ranging from desktop applications to immersive applications. The modelling of bodily cues is more appropriate to full-body virtual humans in immersive virtual environments.

Although the findings from the experiment on kinesics are promising, the significant difference between participant responses to a real person and the affective virtual human suggest that there are aspects of body movement that were not captured in the parametric model. The parametric model changed the qualities of body movement across three parameters. There are other qualities of body movement which were not input into the model such as tenseness of muscle movement. This could be a significant parameter to explore in future versions of the existing model. In addition, Paterson et al. (2001) reported that altering the parameters of an affective animation can change the affective meaning attributed to it. However, it appeared that the quality of body movement was more important to Sadness than Anger. The relative importance of body movement to emotional states could be explored through building parametric models for other emotional states in addition to Anger and Sadness.

Anger and Sadness were mainly chosen because it was unlikely that either emotional state shared behavioural cues. In addition, behavioural properties of Anger are easily recognisable. Future studies could focus on extending the behavioural model to include other emotional states, personality aspects, interpersonal relationships, communicative intents and related factors leading eventually to more and more complex psychological states. These studies could include tests to ascertain if the addition of multiple and congruent behavioural cues aid the process of accurately recognising and responding to the virtual human. Planalp et al. (1996) reported that a number of different behavioural cues were used in tandem to judge the emotional state of an individual in the physical world. How far does this concept apply to full-body virtual humans in an IVE? The literature review suggests that it is better to model multiple behavioural cues associated with an affective state instead of just one. In the studies presented in this thesis, each behavioural cue was studied in isolation. In principle, adding a congruent behavioural cue can only increase the impact of participant responses. However, how much more of an impact on participant responses does adding congruent behavioural cues make? For instance, both Bartneck and Reichenbach (2005) and Hess et al. (1997) argued that increasing the physical intensity of an affective facial expression is associated with stronger recognition rates and strong emotional intensity. In this thesis, extreme behavioural cues associated with Anger and Sadness were used since a dimension of emotional intensity was not included in the experimental design. It would be interesting to explore the relationship between the physical intensity of affective bodily expression and the perceived intensity of the underlying emotional state. This relationship could be further explored using congruent behavioural cues. What is the relationship between physical expression and emotional intensity when more than one behavioural cue is used?

A significant issue to take into consideration while attempting to improve the behavioural fidelity of the virtual human is the visual appearance of the virtual human. Findings from the preliminary experiment on avatar fidelity seem to suggest that increasing either the visual realism of the behavioural complexity of the model in an ad hoc fashion is not sufficient to enhance participant responses to it. However, it remains to be seen if this result carries through to all modalities of expression. Although gaze behaviour is subtle, it is a very important social cue in daily interaction. It will be interesting to see if the interaction effect obtained in gaze behaviour occurs in a similar study on behavioural cues with a different set of primary functions such as kinesics.

The experiments on posture and kinesics focused on ensuring that the participants were not in a direct confrontation with the virtual human. A more direct version of the experiment on kinesics was piloted but could not conducted due to the perceived expression of the mask used in the study. The mask was used in order to ensure that the effects detected in the participant responses were due to the body movement alone. However, a confrontational scenario with the virtual human and an actor could be an interesting way of studying participant responses to a direct interaction with the virtual human. Would participant responses to the virtual human be similar to their responses to the actor?

This research has successfully utilised a combination of objective and subjective measures to investigate research questions dealing with the creation of expression in virtual humans. The results of the analysis suggest some correlation between the physiological arousal and reported copresence. The combination of these two methods to analyse confirmatory type investigations into social contexts in IVEs could prove valuable. In the past, presence questionnaires along with post-experimental interviews have been used to evaluate the effectiveness of virtual humans. However, this form of evaluation is unreliable. Objective means of collecting participant responses could be used in tandem with the traditional subjective means in order to strengthen findings. The use of multiple participant responses to evaluate interactions with virtual humans has the potential to give a more concrete understanding of the interaction between participants and virtual humans. Furthermore, in the case of an exploratory investigation, the additional use of semi-structured interviews can be a powerful way in which to investigate participant responses to virtual humans. The value of qualitative research as argued in Section 3.1 was justified in both the second preliminary experiment presented in Chapter 4 and the main experiment presented in Chapter 5. For instance, the qualitative research was vital in uncovering participants' interpretation of the social scenario around the agents in the experiment on posture (Chapter 5).

Despite the complexities involved, virtual humans present promising avenues for research into social behaviour. Virtual humans are an important and powerful part of content in virtual environments especially with respect to invoking participant responses. In evaluating these responses, it is important to take into account the differences due to individual traits. This is a diverse area of research and this diversity of research is itself a challenge to researchers in this field: each resulting virtual human system has been designed to investigate a particular aspect of non-verbal communication. This thesis focused on building parametric models of two behavioural cues: gaze behaviour in dyadic communication and affective kinesics. Future work will building on these models by adding other modalities of expression and investigating the impact of increasing behavioural fidelity in expressive virtual humans on a variety of participant responses.

Appendix A

Published papers

1. (Mortensen et al., 2002): Mortensen, J., Vinayagamoorthy, V., Slater, M., Steed, A., Whitton, M. C., and Lok, B. (2002). *Collaboration in tele-immersive environments.* In the Eight Eurographics Workshop on Virtual Environments.

2. (Freeman et al., 2003): Freeman, D., Slater, M., Bebbington, P., Garety, P. A., Kuipers, E., Fowler, D., Met, A., Read, C. M., Jordan, J., and Vinayagamoorthy, V. (2003). *Can virtual reality be used to investigate persecutory ideation*? Journal of Nervous and Mental Disease, 191(8).

3. (Garau et al., 2003): Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., and Sasse, A. M. (2003). *The impact of avatar realism and eye gaze control on the perceived quality of communication in a shared immersive virtual environment.* In SIGCHI.

4. (Vinayagamoorthy et al., 2004b): Vinayagamoorthy, V., Garau, M., Steed, A., and Slater, M. (2004). *An eye gaze model for dyadic interaction in an immersive virtual environment: Practice and experience.* Computer Graphics Forum, 23(1).

5. (Vinayagamoorthy et al., 2004a): Vinayagamoorthy, V., Brogni, A., Gillies, M., Slater, M., and Steed, A. (2004). *An investigation of presence response across variations in visual realism.* In the Seventh Annual International Presence Workshop.

6. (Steed et al., 2005): Steed, A., Brogni, A., and Vinayagamoorthy, V. (2005). *Using breaks in presence to identify usability issues.* In the International Conference on Human Computer Interaction, Volume 11.

7. (Vinayagamoorthy et al., 2005): Vinayagamoorthy, V., Steed, A., and Slater, M. (2005). *Building characters: Lessons drawn from virtual environments.* In Towards Social Mechanisms of Android Science: A CogSci 2005 - XXVII Annual Conference of the Cognitive Science Society workshop.

8. (Gillies et al., 2005): Gillies, M., Vinayagamoorthy, V., Robertson, D., and Steed, A. (2005). *A platform independent architecture for virtual characters and avatars.* In Intelligent Virtual Agents, Poster presentation.

9. (Vinayagamoorthy et al., 2006a): Vinayagamoorthy, V., Brogni, A., Steed, A., and Slater, M. (2006). *The influence of a virtual character posture model on reported presence in an immersive virtual environment.* In the ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications - VRCIA.

10. (Edlinger et al., 2006): Edlinger, G., Guger, C., Leeb, R., Friedman, D., Vinayagamoorthy, V. and Slater, M. (2005). *Controlling Virtual Environments by Thoughts.* In the Fifth Forum of European Neurosciences, Poster presentation in the Cortex and Thalamus session.

11. (Vinayagamoorthy et al., 2006b): Vinayagamoorthy, V, Gillies, Marco, Steed, A, Tanguy, E, Pan, X, Loscos, C and Slater, M. (2006). *Building Expression into Virtual Characters.* In Eurographics 2006 - State of the Art Report (STAR).

12. (Brogni et al., 2006): Brogni, A., Vinayagamoorthy, V., Steed, Anthony and Slater, Mel. (2006). *Variations in Physiological Responses of Participants During Different Stages of an Immersive Virtual Environment Experiment.* In the ACM symposium on Virtual Reality Software and Technology (VRST).

Appendix B

Kinesics: The Case of Clever Hans

One of the earliest known studies into body language is the case of Mr. Von-Osten's horse: '*Clever Hans*'. Clever Hans and his owner enjoyed world wide acclaim for the horse's ability to do arithmetic and other problem-solving tasks. The horse, Hans, had been trained by a Mr. Van Osten to tap out the answers to questions with its hoof.

Oskar Pfungst demonstrated that the horse's claimed abilities were due to an artifact in the research methodology. Pfungst found that the horse could get the correct answer even if Van Osten himself did not ask the questions, ruling out the possibility of fraud. Pfungst devised a series of studies which determined unequivocally that Hans was successful only when there was visual access to the questioner. Han's accuracy declined as distance to the questioner increased and was markedly reduced by blinders. In cases in which the questioner was ignorant of the answer, Hans failed miserably. Pfungst (1911) observed that as Hans got closer to the correct response, the questioner's tension level would increase and at the moment the horse approached the correct *tap*, tension release would yield a tiny, involuntary movement in the questioner prompting Clever Hans to stop tapping.

Han's intellectual achievements were due, not to independent thinking, but to the recognition of behavioural cues provided by unconscious involuntary movements in the facial expression and posture of the questioner (Pfungst, 1911).

Appendix C

Preliminary Experiment on visual realism

C.1 Code Lists

Code-Filter: All

HU: Street
File: [D:\Street\Interviews\Street.hpr5]

AAmerican	**AGenderless**	**AMales**
AOld	**AppearanceReal**	**ASaudi**
AShort	**ATall**	**AWearingRobe**
AWhite	**BIPS**	**ComputerDriven**
EAAInteractions	**EAcknowledgement**	**EEyeContact**
EInteractivity	**ENoInteractivity**	**ENoTalk**
ENoVocalApology	**EResponsiveness**	**ESmiles**
EStop	**EVisualRealism**	**EWalkThrough**
ExperienceComputer	**FactorCartoon**	**FactorFemale**
FactorMale	**FactorNonRepeat**	**FactorPhoto**
FactorRepeat	**NoAudio**	**NoCD**
PAAnnoyed	**PANotWantContact**	**PArrogant**
PARTICIPANT Expectations	**PCalm**	**PCold**
PERCEIVED Realism	**PERCEIVED States**	**PEyeContact**
PGrumpy	**PLookedUnreal**	**PNeutral**
PNoAAInteractions	**PNoAcknowledge**	**PNoBodyLanguage**
PNoChangesStates	**PNoEmotion**	**PNoEvasiveAction**
PNoEyeContact	**PNoGazeBehaviour**	**PNonResponsive**
PNoSmile	**PNoVocalApology**	**PPreoccupied**
PPurposefull	**PRandWalk**	**Predetermine**
Programmed	**PSideStepped**	**PUnAware**
PUnfriendly	**PUnrealMovement**	**PWalkedThroughMe**
PWalkThroughLamppost	**RAnnoyed**	**RAnxious**
RAnxiousConverse	**RApologised**	**RAvoidBump**
RAvoidInteracting	**RDidNotInteract**	**RDidNotNeedInteract**
REALISTIC Responses	**RFeltBehUnImp**	**RFeltOutsider**
RFeltUnImp	**RFeltUnreal**	**RHesitant**
RIdentifiedWith	**RIgnored**	**RLostInterest**
RMemory	**RNotIndentify**	**RNotWantInteract**
RScared	**RSmile**	**RWantedHelp**
RWeird	**RWorry**	**SocialContext**
StrollingBehaviour	**TApproach**	**TBump**
TInteract	**TreatedAsAgent**	**TStand**
TStare	**TStop**	**TTouch**
TWalk	**UESmile**	**UnrealBehave**
WalkingAsExpected	**WalkingReal**	**WalkingUnReal**
WEIRD Attributes		

Figure C.1: Code list: second preliminary experiment

C.2 Sample Interview

Experiment: Realism and Believability 2004
Environment: Street
Project: EQUATOR

DATE	[Thursday 22nd Jan]
TIME	[13.30]
PARTICIPANT ID	**[2][1][1][1][1]**
EXPERIMENTERS (2)	[Vinoba, Andrea]
TAPE NUMBER	[02 - 03]
SIDE (A OR B)	[B - A]
COMMENTS	[Any Comments: inaudible experimenter? Participant?]

[START OF INTERVIEW]

E Participant's id code 21111. Ok, can you describe what you saw... the street. What do you think about the street?

P Its very long. Uhhh initially very dull, bleak, no trees, no cars, uhmm lots of shops. And then uhmm people, not very people, sparsely populated, so in that way its very strange.Its uhmm... its...

E Uh hum. And what your feeling to the street? I mean, uhhh can you describe some emotion that you get uhhh...

P I thought... well while I was going, oh this is a really dull street, so I wanted to go to the end to see where it would take me and all the buildings see the same. And then I thought, ok lets explore the shops, and then I went into the shops... I got... I like going and finding where this was going and I presumed what they were selling and where I could go next and in coming out whether there are many people around, I better look and I looked and I bumped into a woman. Uhmm and then the darkness and then was coming out.

E Uhmm. What kind of shop did you see?

P Uhh the... mainly clothes, there was precision and waterstones. Waterstones, I remember. Uhmm books... but I wasn't really... unless it was like the books that I was really interested because that is my field. I was Intrigued more of where I could go rather than the items in the shop. So, I wasn't really looking at the items. I wasn't... there was a shop, I was there to find out whether I could through the shops. What was next...

E Did you... Did you recognise the place as a place where you have been before?

P It would probably have been, Saudi Arabia. I have been in Saudi Arabia... Saudi Arabia... yes, the emptiness of the place and the building. There were people there and I thought they were saudis'.

E But it's a place where you could have been?

P Yes, yes I could say... yeah but no I don't recall but similar to... in reality to shopping centres in Arabs, yeah.

E Uhmm...

P On a Friday morning

E Did you notice something around, the background? You know sky or... the horizon... the horizon.

P No, I was so....I won't...My, my intention was the end because I looked...and there were no trees or anything and then I thought this was odd and I wanted to see what was at the end of the street. To go up to it and get there and I realise that there is more shops and I never really looked... bothered about the arising at all. Except I think, once. And then I realise that it was...I saw the top... dark...and then came back.

E Ahh. Ok. Uhmm So you...

P It was, it was sunny, very bright. Just isn't real. It was a clear sunny day, blue sky.

E You mentioned that you uhmm attributed a nationality to the avatar.

P Yes

E You thought they were Saudis.

P Yes

E Uhmm what made you think that?

P Uhhh he was wearing a throbe, a traditional dress... the white

E Oh the white... white

P Yeah..

E Ok, And uhmm how many of them did you see?

P Uhmm, the Saudis...I saw...I think I saw one where I thought that they are in Saudi, and then I saw again... another one and then I bumped into one. Another lady there as well. She wasn't Saudi. She she reminded me in American in green.

E How many people, how many people do you think there was actually?

P No more than ten.

E And uhmm which kind of people apart from this Saudi.

P Now that I think about it, it was like if I was there because uhmmm... this lady I mentioned, who was in green and I thought she was American because they were those kind of clothes. Nothing wrong with that... But then they were...I was so interested in the shops and and the places...I wasn't even focused even in the distance. And they were wearing white... so I thought they were Saudis. Uhmm...

E What do you think they were doing there?

P The...going about their business. There weren't any shoppers. There weren't people in the shops. So that is why I thought I was a very quiet Friday morning... when there is prayer. And they were going to another place. I doubt if they were using the street as a... they weren't window shopping. They weren't. And most of them I think were all singles. So they are on their own... so... they were determinally going on by their business. Going somewhere. And then... yeah... and then it was funny, in the end when the light went down, and I had this fear of...I had to get out. Trying to find the way and I was lost. And then I had this...i saw a woman go by, and I wanted to stop her but I thought she was running as well and I had this impression... feeling that everybody was doing the same as I was. That they were speeded up. And that they were hovering at the end of the day. They had to get somewhere. And then I heard your footsteps...

E No no no its ok. Did the feeling remind you of any...any other experiences you have had before? Of being lost or being in the dark?

P I...I....when I was a child, I was scared of the dark, yeah. The....but in being lost...unless it is something to do with movies I have seen. Horror movies. In a strange place and you have to...had this feeling when I was there. Trapped...and I had to get out.

E Are you normally hesitant to be in the dark? Or...

P I would uhh...not normally, no. But it can...but not at not at night in the street or anything. Am not hesitant like that. Normally in a very strange place. And I think the suddenness of it suddenly came down in the night...

E So was there...

P The darkness...

E Was too fast...so...I mean

P It was like a...uh...it was fast but it is just as if you...again...sudden storm comes over. That's was what I had the feeling because I had been storms and suddenly black. And it was as if a sudden storm had come.

E Did you mention...did you have this transition back.. but you mentioned you didn't have a transition back there. Or did you?

P I did and uhhh...for about a second. Because it went dark and I had a certain job and then I remembered what you said. That that that is the time to get out and I was like...but then I was hesitant and I went back again and I was in the place and I had to get out and it wasn't you know, get out to the lab. Get out of the place. Find the door, that was it, find the door.

E And the sight of seeing another person there?

P Sorry

E Uhmm. You saw the lady...

P I did, yes

E Standing by and you wanted to ask for directions?

P Yes but she was gone to quickly.

E Ok

P And then I started speeding up...I was rushing...I was rushing...a bit more.

E How many times did you try to interact with the people there?

P Just that once really. I think when I bumped into the lady, I returned a smile and I said give my apologies. I didn't say anything but you know how you smile at people. Uhmm but that was it. And then I got...wanted to stop her.

E Did she respond?

P No, no. She was hurrying on her way. She was...she was in a hurry to go.

E How did that make you feel?

P It made me more worried.

E Is that your normal reaction in a normal street? If... away... If you sort of bumped into a person, you apologised and they didn't pay attention, would that be your normal response? Or was it something new?

P This was new. This was a new... I think because the element of fear was there. I wanted contact, I wanted help. But when you are lonely in a street, you are talk to someone and you don't help. Its... you know that's busy life. Its ok. But uhh... yeah

E So, you got lost? So but you have an idea more or less of the layout of the street.

P I thought I did initially. Now I got more panicky. So I went... I thought it was just one long street, so I came out of the shop and tried to go down... left... thinking the door. But the door wasn't there. And then, then I got really lost. And it really seemed to have more sparse. It was more complicated than I imagined than I realised I hadn't. Because when I came in, when I was in the starting street and then I was in the street. I walked, thought I remembered the way I am going that I am just going down one true path and shops... was a branch and when afterwards, I realised I had gone through many because I had gone the shops and come out through the door. And I got confused. And I couldn't find my way. So...

E If I asked you to draw a layout... a layout of the street, more or less. So start from when you came out and you know, the door opening and what happened.

P Oh, in the very beginning. Yeah. It was really just straight. Through the door and then there was like a... that was it. And like that. And there were streets. These were streets. And here they just had buildings. And all this buildings yeah. I came in through there.

E Uhh. Instances that you bumped into the lady or when you had transitions in the lab. Do you recall around which part of the street you had these transitions?

P Uhmmm. The lab as I said was at a very beginning. Because of the expectations and thinking it was a boring street. And then I... it was... when I turned to look at the buildings very closely at the beginning. I saw the corner of the wall. And then I knew I was in the lab. That was very obvious to me. And then I forgot really, uhmm and then the lady I bumped into was here. I was going cross the street to waterstones.

E Ok, could you just write waterstones...

P And it was... I remember now, it was uhmm a lady... it was a... it was cobalt... was that... these were cobalt stones. That was new when I enter. So, I think that gave... that gave me the impression... that I was in a different place from the beginning. It was just white. I was crossing cobalt

E When you get the transition the beginning, how long did it take to recover. Come back in... the street?

P Fairly quickly. It was... you press... when you press the button, it stopped, the frame stops and starts again and then almost as soon as it starts I am back in the street.

E And you say, you had this two or three times you remember. And...

P Probably more...

E You think was similar time, the recovery... recovery time or the intensity of the transition was the same in all of them or changed?

P The... I think they are all the same, the intensity or the length. I don't really remember later on coming back to the lab, uhmmm I think uhmmm it was vague. If I do... towards the end... in the middle I press that... it was very vague. I was back in the lab. In the beginning it was very clear. And then it was like vague. It was much quicker, I was back into reality. Wasn't this marked transition.

E So, some people stood out to you?

P Yes

E One was the lady you bumped into?

P Yes

E The other one would be the Saudi man?

P Yes

E Uhmm, who else that you recall?

P Well, it was only those three because they were all the same...

E Who was the third?

P Who... the third... the third was the one that was walking along the street and I bumped into him. And there was a lady in green who I bumped into. I actually hit and apologised. Watched her and she was there... and then there was this one... she was in white and grey and she was dashing down that way.

E Dashing towards the end of the street?

P Yeah. And there were few others but they didn't really... didn't impact... they were just figures.

E Ok.

[Cassette change]

E Ok

P Can you repeat the question?

E Lets see if I remember the question. Uhmm, we were talking about the number of people who stood out.

P Right

E Uhmm, if you, if you feel some emotional... emotion... attribute. Personality to these people? In general.

P The... the Saudis were walking... just going about their business quiet calmly. And... then the lady I felt was a bit annoyed when I bumped into her and then the lady who was a... when I I tried to stop at the end was also one or two. There were all, all rushing. And they were in a hurry and they were they were in the same situation as I was. I got that. But they knew where they were going. Maybe now, that I think about it. Maybe a rain was coming down, it was some feeling or bolting... they wanted to get away just as much as I did. But they knew the way, I didn't.

E So this was after you, after it went dark?

P Yes

E This feeling of wanting to hurry out

P Yeah, yeah otherwise they were just going about with no frantic case. They weren't window shopping or anything, just...

E You say, the first question ask you.. stood there... say that or whole experience, you are in the street and part of... the beginning, the very beginning.

P Yes

E Uhh so ask you to draw a graph to here... time... from here where we open the door, that the experience, and here you call presence is the sense of being in the street.

P Right

E If you can draw, if you had a... the intensity during the time.

P I am not sure, I don't think I understand quiet what I have to draw...

E If you feel a sense of being in the street...

P Yes

E You say you are very high...

P Yes

E So, can you draw if change during the whole experience from the beginning when we open the door to... to the end... if you had something. This is your sense of being there. Say when you came out of the training room, you said you felt the sense of being in the lab more because you uhh noticed the corners?

P Right

E So, I guess that would be a low sense of presence. Here the street...

P Oh, oh right, that's the beginning?

E Yeah,

P So, high

E Well, that's the virtual place or rather the transition... so what we want to know is when you transited to the lab and when you came back to the street over time. This is the street, this is the lab... where you are

P Oh ok, this is this is the time, this is the beginning and this is the end?

E Yeah

P So, if at the beginning I was in the lab, so its here. Right. And and then it was really...I don't... how do we realise going out into the street

E You know, going up would be a line... continuous line... was it like immediate or... I don't know, you can do. It was very fast and then here or slow.

P The length of the time compared to... yeah

E Say this is the end of the experience.

P Yeah

E So you are coming out of the lab. Ok, you have reached the door.

P Yeah

E Uhmmm, did you go into the... did you rec... or did you become part of the street already or

P No

E Ok, so I suppose it would go slowly, uhmm..

P Oh, I see and then I said I was in the street almost all the time...

E So you kind of come up

P All the time that I was there...

E And then, then did it increase throughout the time you were there or did it reach a steady state or?

P It increased

E It increased.

P Well, maybe one... maybe a bit flatter...

E Ok

P But then at the end it really increased.

E Oh ok, so this was after it got dark?

P Oh no and during when I bumped into the thing and it made me... and I was... I bumped into it

E Ok, then you bumped into it and what happened? Did you get the transition back into the lab or did it increase your sense of being there?

P It increased my sense of being there. I bumped into this... door...

E Ok

P Yeah and then it went straight right to the end

E So this would be the end of your experience? I guess

P No sorry, it last all like that to to the end. Until you come in with your steps.

E Ok, right. Uhmm. What was the other question. Hmm. It reminded you of Saudi, you said, the open space. The shopping street on a friday morning.

P Yeah

E But you could attribute hmmm a place....in real life... real world that...

P Very much so. Very much so. Yes. Very much so.

E Ok. Is there anything else you want to add, however much insignificant? Something that stand out? That stood out in your experience?

P The physical feeling... the emotion... the and....when you bumping in...you are looking, you get, you get dizzy. I got dizzy because I was....but it wasn't like I...It was like if I was very active. Though I was dizzy from being active. When you are turning around like this, you get dizzy. Uhmmm that was it. In the beginning it was a video, a video archid. And then definitely

E You, are you surprised by your overall reaction?

P Very, very, very…I…I…because at the beginning, I thought well I keep seeing these corners of the walls, so I am going to keep seeing them for the time the…the…and then then the beginning I was thinking well there was a wall in front of me so, I have to be careful, I can't injure myself because I thought I got to move but then, then I was moving but not moving.

E Yeah

P I wasn't aware of physically being there and having to physically move. I felt myself move…I lost the sense of…

E So, did you feel a sense of not being able to move slowly at the beginning of the experience or during your encounter…during the experience at some point?

P Slowly at a…I wanted to get to the end. But it was right at the beginning. It was a boring street and I wanted to see what was at the end. And I was pressing the thing. And the street didn't come…because…first few seconds felt it was coming towards me and then I got into the street and it was and it was normal speed. Just travelling as I normally do and then at the end, it was…we were all rushing to go somewhere.

E How about the other people around you? How did they compare with you? In terms of going somewhere.

P Perfectly natural. Walking about their business. Not agitated except the end…and naturally they were…yeah same as me definitely.

E Uhh was the street something that you saw? Or something. Somewhere you visited.

P Well, I was there, I visited. I was…It was…

E Uhh…you know that you can't talk about this for a few months because we are running experiment.

P Uh hum. Yeah.

E If you have to explain to your friends…

P Right.

E What is happening this afternoon. How you describe it?

P I would say the beginning it was a video arcade and the street I saw and and the street you saw was like being in a video arcade. And then I was in the place and it became a place for me because I…I was in Saudi Arabia. And it was place yeah. And then when you came, it was back in the lab and it was an experiment.

E Ok, thank you very much.

P Not at all.

… … …

[END OF INTERVIEW]

C.3 Code Descriptions

AAmerican	"Character was American"
AGenderless	"Participant couldn't tell or remember gender"
AMales	"characters were males"
AOld	"Character was old"
AppearanceReal	"One participant thought the higher realism character looked real"
ASaudi	"characters were Saudis"
AShort	"Characters were short"
ATall	"Characters were tall"
AWearingRobe	"Character was wearing a robe"
AWhite	"Characters were white"
BIPS	"Breaks in Presence caused by Characters."
ComputerDriven	"Characters were computer driven"
EAAInteractions	"Participant expected characters to interact with each other"
EAcknowledgement	"Participant expected characters to acknowledge them"
EEyeContact	"Participant expected character to turn and look as acknowledgement"
EInteractivity	"Participant expected characters to interact"
ENoInteractivity	"Participant did not expect that character could interact"
ENoTalk	"Participant did not expect character could talk"
ENoVocalApology	"Participant wasn't expecting character to say excuse me"
EResponsiveness	"Participant expected character to respond to them"
ESmiles	"Expected character to smile"
EStop	"Participant expected character to stop in response to participant being in the way"
EVisualRealism	"Participant expected character to look more real"
EWalkThrough	"Participant expected character to walk through"
ExperienceComputer	"Participants' experience from computer interfaces as a reason for expectations"
FactorCartoon	"All interviews in which characters were cartoonish"
FactorFemale	"All participants were females"
FactorMale	"All interviews were male participants"
FactorNonRepeat	"All participants were in the non-repeatitive conditions"
FactorPhoto	"All participant were in the high visually realistic conditions"
FactorRepeat	"All interviews are with repeated textures"
NoAudio	"No Audio"
NoCD	"No Collision Detection"
PAAnnoyed	"Character was annoyed"
PANotWantContact	"Agents did not want contact"
PArrogant	"Characters were arrogant"
PARTICIPANT Expectations	"Expectations of Participants"
PCalm	"Characters were calm"
PCold	"Characters were cold"
PERCEIVED Realism	"Perceived realism of character"
PERCEIVED States	"Perceived States of character"
PEyeContact	"The character looked at me!"
PGrumpy	"Characters seemed grumpy"
PLookedUnreal	"Characters did not look real"
PNeutral	"Agents were of neutral attitude"
PNoAAInteractions	"Characters did not interact with each other"
PNoAcknowledge	"Characters did not notice or acknowledge me"
PNoBodyLanguage	"Characters did not have appropriate body language not even head movement"
PNoChangesStates	"No changes in perceived state"
PNoEmotion	"Characters had no emotional aspects"
PNoEvasiveAction	"Character did not move out of the way (side step) or stop."
PNoEyeContact	"Characters did not look at me"
PNoGazeBehaviour	"Characters gaze behaviour was not real"
PNonResponsive	"Characters did not respond"
PNoSmile	"Characters did not smile"
PNoVocalApology	"Characters did not say Sorry or Excuse me"
PPreoccupied	"Characters perceived to be preoccupied with their business"
PPurposefull	"Characters were busy, purposeful, and goal-oriented"
PRandWalk	"Characters were randomly walking (Behaviour)"
Predetermine	"Character behaviour looks as if it were predetermined."
Programmed	"Agent is programmed to avoid me"
PSideStepped	"Agents side stepped the participant"
PUnAware	"Characters did not know I was there"
PUnfriendly	"Characters were unfriendly"
PUnrealMovement	"Movement was not real"
PWalkedThroughMe	"Characters walked through me or did not avoid me"
PWalkThroughLamppost	"Characters walked through lamp post"

RAnnoyed	"Participant got annoyed"
RAnxious	"Participant was anxious of a possible negative reaction"
RAnxiousConverse	"Participant was anxious that they won't have a great conversation"
RApologised	"Participant apologised to the character"
RAvoidBump	"Participant avoided bumping into the agents"
RAvoidInteracting	"Participant avoided interacting with the character"
RDidNotInteract	"Participant did not interact"
RDidNotNeedInteract	"Participant did not need to interact with them"
REALISTIC Responses	"Responses observed to characters"
RFeltBehUnImp	"Participant felt their behaviour did not matter"
RFeltOutsider	"Participant felt an outsider or invisible"
RFeltUnImp	"Participant felt unimportant"
RFeltUnreal	"Characters did not feel real"
RHesitant	"Participant hesitated to walk into character"
RIdentifiedWith	"Characters and I were in the same situation"
RIgnored	"Participant ignored the agents"
RLostInterest	"Participant lost interest in characters"
RMemory	"Reminded participant of a previous occasion"
RNotIndentify	"Participant didn't have same interests as characters"
RNotWantInteract	"Participant did not want to interact"
RScared	"Participant felt scared of character approach and of lack of help"
RSmile	"Participant smiled at character"
RWantedHelp	"Participant wanted help/contact to ask for directions from the agents"
RWeird	"Participant turned and looked at character's backside"
RWorry	"Participant felt more worried"
SocialContext	"The perceived social context"
StrollingBehaviour	"Agents were not strolling or looking in shops"
TApproach	"Participant tried to approach the character"
TBump	"Participant tried to walk through character or bump into them"
TInteract	"Participant tried to interact"
TreatedAsAgent	"Participant did not treat agents as real"
TStand	"Participant tried to stand in front of character"
TStare	"Participant tried to stare at them"
TStop	"Participant tried to stop character"
TTouch	"Participant tried to touch character"
TWalk	"Participant tried to walk with them"
UESmile	"Unexpected smiling behaviour"
UnrealBehave	"The character behaviour did not look real either due to body language, gaze behaviour or jerky movement."
WalkingAsExpected	"Walking behaviour was in keeping with expectation"
WalkingReal	"Walking looked real"
WalkingUnReal	"Walking was not real"
WEIRD Attributes	"Attributes that are unexplained"

C.4 Network of Codes

This section contains the images generated to give a Visual Abstraction of the qualitative analysis of interview data collected in the preliminary experiment on visual realism.

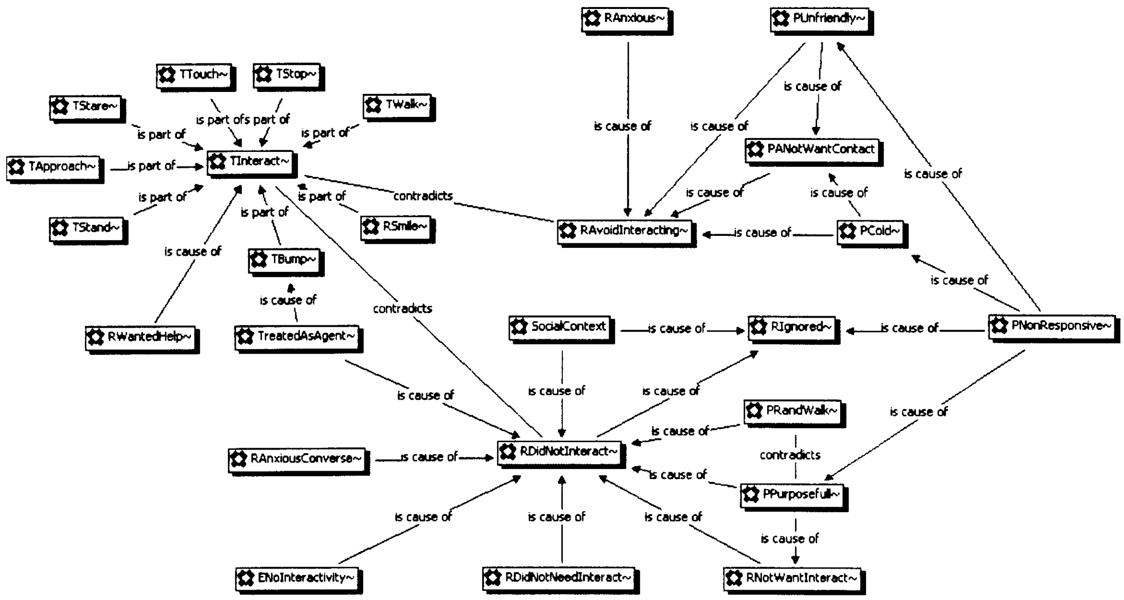


Figure C.2: Participant interactions with the Agents

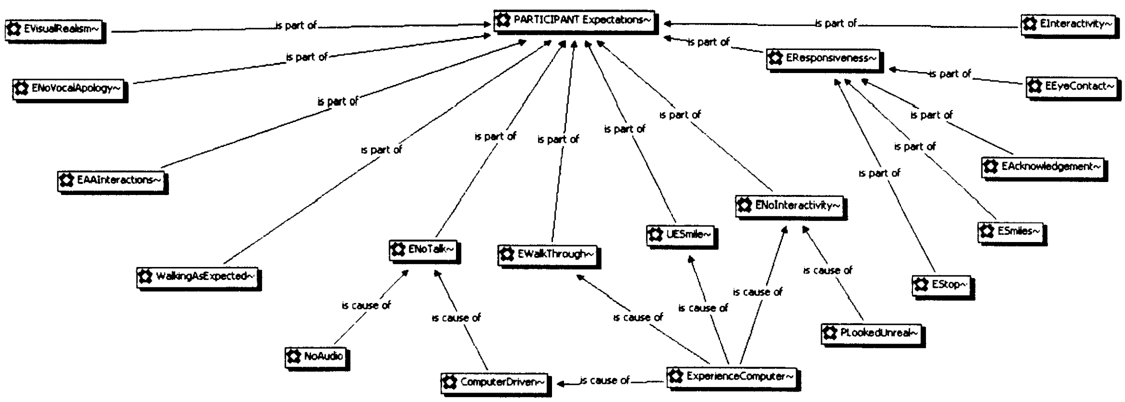


Figure C.3: Participant expectations of the Agents

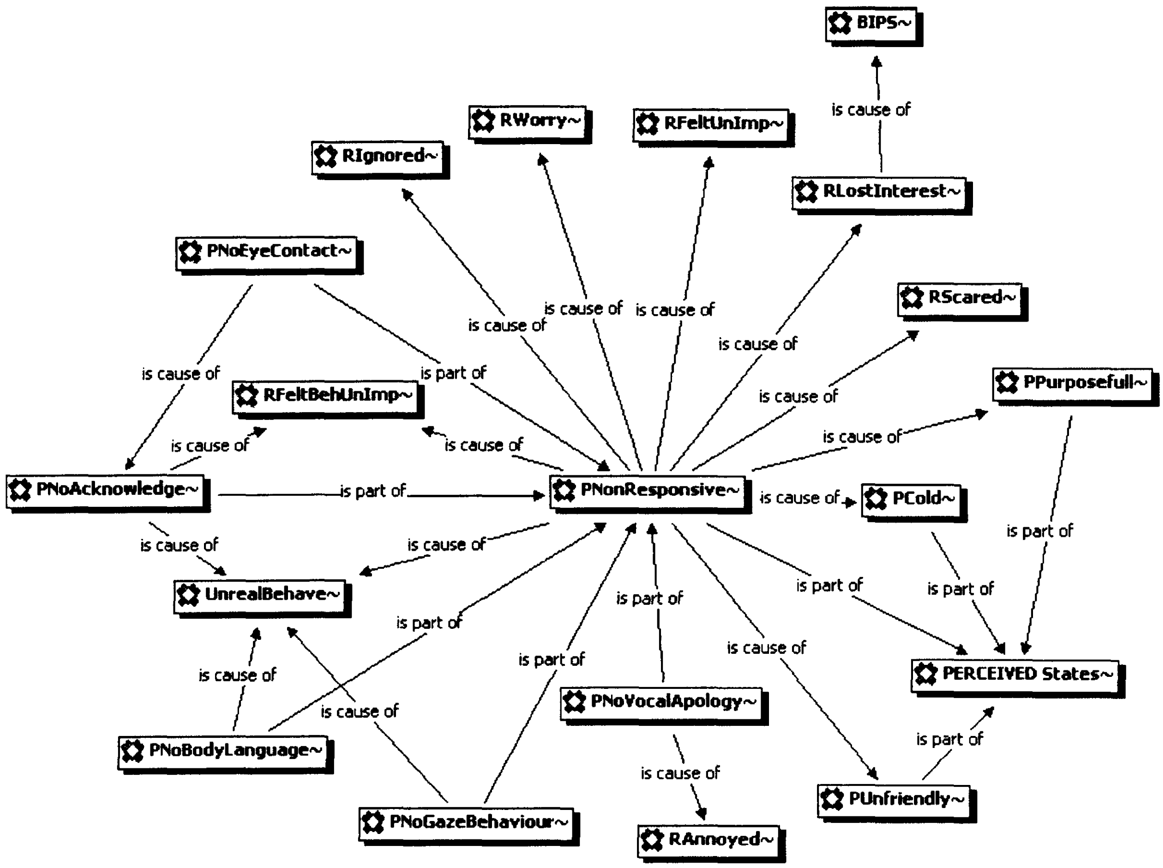


Figure C.4: Causes of perceived non-responsiveness in Agents

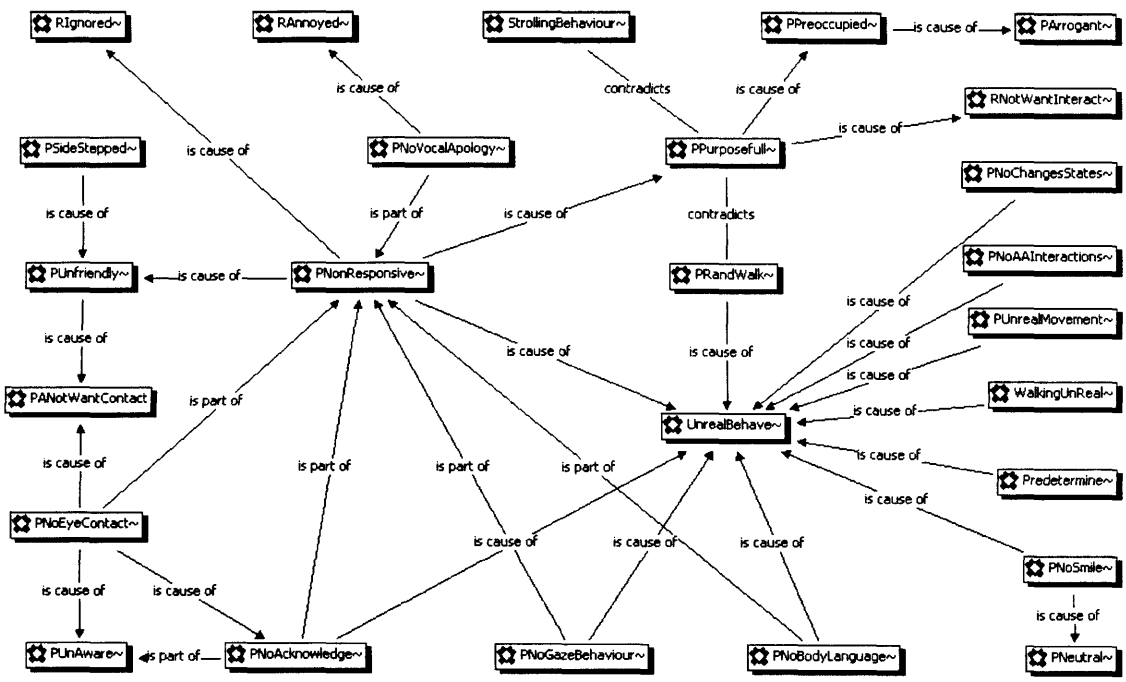


Figure C.5: Agent Behaviours participants did not view as realistic

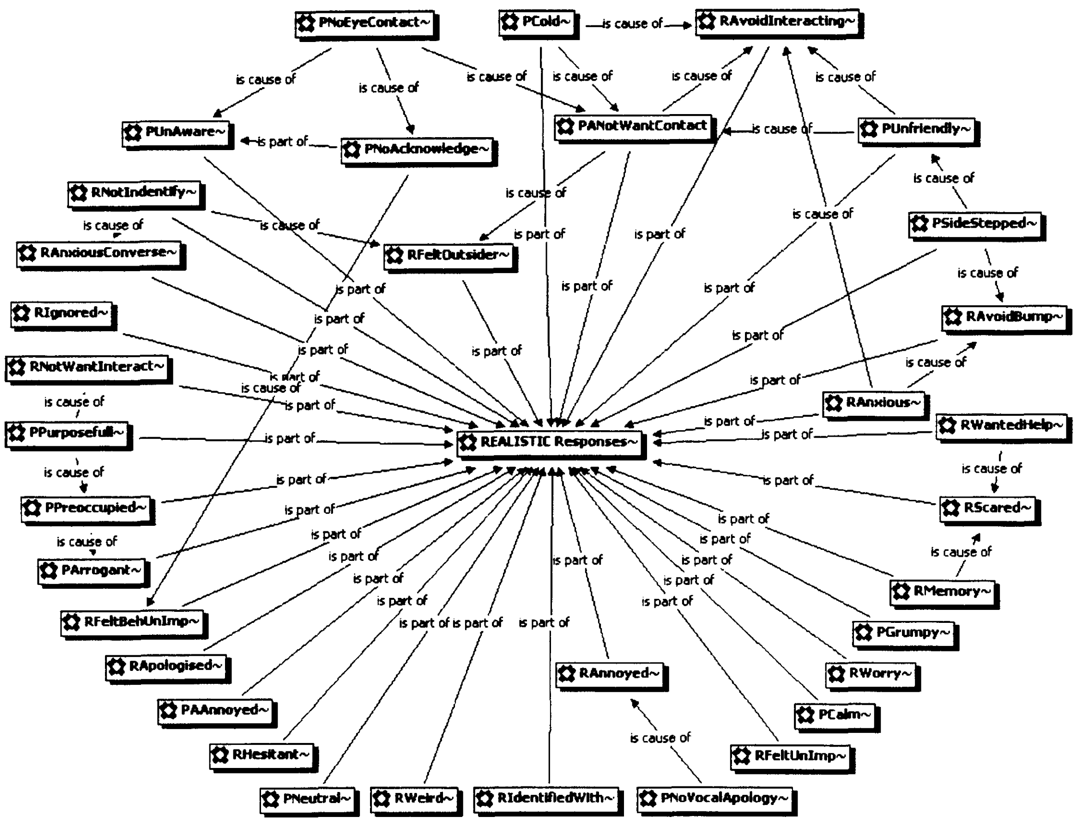


Figure C.6: Realistic participant responses to Agents

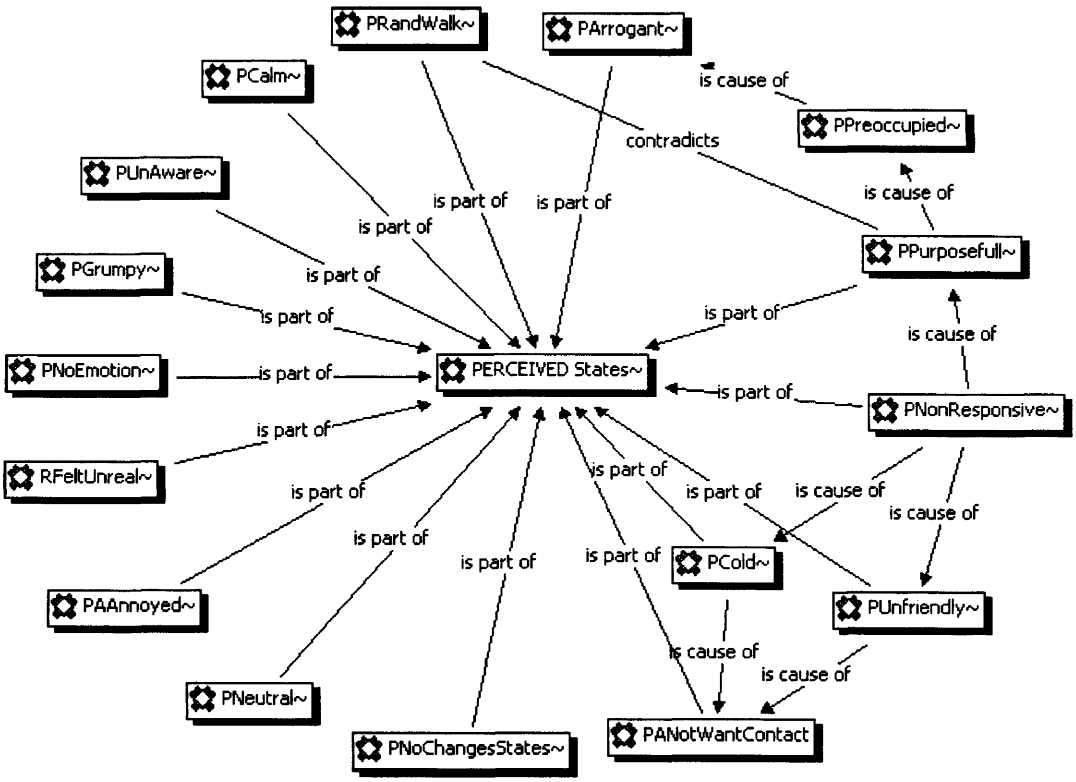


Figure C.7: Perceived Agent attributes

C.5 Participant's findings from interviews

Generally participants reported that the agents lacked an emotional dimension, realistic gaze behaviour, gestures, and natural body movement: "*Some people, their body posture might change slightly, if they acknowledge somebody. I cant really explain it. I am not an expert butI just noticed these things.* (P31121)". Participants also indicated that a lack of agent-agent interaction in the virtual street would have added to the behaviour fidelity of the agents: "*You expect people not to do the same action like walking ahead. I mean they could stop at the window of a shop, chat with each other or smile probably* (P31221)". However, the findings presented in the following section focuses on participant-agent interactions and participant responses.

C.5.1 Participant-Agent interactions and expectations

Participants reported trying to interact with the agents by first approaching the agents, standing in front of them and trying to stop them or bump into them with the expectation that the agent would respond or interact: "*I thought that maybe the fact that I just approach them would make them feel... would make them have a reaction and I saw no reactions* (P21221)". These were some of the expected forms of interactions anticipated prior to running the experiments, however, some participants reported expecting responses to more subtle forms of interactions such as staring or looking at the agent. "*I should have... I should have tried to talk to people or to touch them but what made me not to do that was the fact that every time I look at them, I had no reply* (P21221)". Most participants tried to interact with the agents in order to see how their actions would cause the agents to react: "*I don't know what to expect, I was just hoping that they did sort of, there did be some sort of interaction like, they might like stick their hand out or like smile at me or something* (P31121)". However, some participants tried to interact to get information, for instance, one participant got "lost" in the virtual street and wanted to ask for directions: "*I wanted contact, I wanted help.* (P21111)". Whatever the form of interaction, the participant had preconceived expectations of how the agent would react.

Most participants expected the agents to respond to their presence using a number of behavioural cues including eye gaze, facial expressions and bodily cues: "*If you see a person, you might like, it will show like on your face, you might like acknowledge that. You might, if you are a guy and you are walking on the street and you look at a good looking girl, you might smile or something, right? Or if you see like a tramp or something, you might have a mournful look on your face, you feel sorry for him or something, so its just expression of feedback* (P31121)". This was especially the case when the participant made an attempt to interact with the agent: "*I mean it was easier to get into the environment like the buildings and things because buildings don't tend to do much, they just like stand there and they exist. Whereas people, you tend to expect to do a bit more really... like stopping when I move near them, in front of them, maybe if I walk around them, they might turn and look at me. And sort of yeah, just like tending to acknowledge my existence rather than sort of looking as if they were walking in a predetermined way* (P21121)".

Some participants did not even attempt to interact with the agents because they did not expect the agents to respond either because the agents looked computerised: "*They didn't seem like people enough to try. I didn't smile at them or I just tried to get in their way to see what they will do really. And I didn't* *really expect them, from their look, I didn't expect them to do anything else.* (P52111)" or because of their prior experiences: "*I just stopped and stared, I mean, I guess I, from what I know about computers interfaces, I didn't really think that talking to them or waving at them would really attract their attention but that's just giving away my experiences with machines* (P11211)". Some participants did not interact with the agents because they did not need to: "*Usually I don't interact with... you know, if I am walking in a street, its rarely that I stop someone unless I need information. And since I was there to explore, you know I just wasn't feeling the need of information, I just didn't need to interact with them* (P22111)" or because of the social context provided to the participant in the virtual street: "*If they are just normal people walking on a street, I wouldn't try to interact with them.* (P41111)". One participant did not interact with the agents for fear of a negative reaction: "*I wasn't sure what their reaction would be to me. I suppose I didn't really want a negative reaction* (P32211)".

C.5.2 Non-responsive Agents: The participant perspective

The most significant missing element pointed out by most participants were the agent's lack of responsiveness and acknowledgement of the participant: "*They were walking in a very mechanical way and they weren't even looking at me and when I bumped into them, they had no feeling, they didn't say anything, they didn't even move their heads* (P22221)". This lead some participants to treat the agents more as virtual objects: "*You see then you may bump but again because it was clearly uhh I don't know how to explain; one felt being invited to bump into people because what else could you have done because there was no reaction... Usually if you bump into people, they react angrily* (P51111)". The lack of responsiveness lead participants to feel that their behaviour in the virtual street had no consequence: "*I got the feeling after being there for a little while but it didn't matter what I did, that my behaviour wasn't having any effect on what was going on in the rest of the simulation* (P21211)" while a lack of acknowledgement caused through a lack of participant-agent eye contact, made participants feel ignored: "*Yeah, it was as if they were ignoring me but like they knew I was there but they were just ignoring me, like they weren't very polite* (P31121)".

Even though, a lack of agent responsiveness resulted in minimal participant-agent interaction, a significant number of participants attributed this to the agents being "purposeful": "*I just thought that they were just busy. You know that they were going somewhere. And I wasn't part of the process. And I didn't find that threatening or annoying or anything, it seemed quite natural* (P32111)"; busy: "*I think they seemed like they were in a rush, they walked quite quickly. They seemed like, they knew where they were going.* (P32211)" or arrogant: "*It did kind of seem like they were kind of stuck on themselves. You know, there are too good to talk with others around them* (P12221)". A lack of responsiveness led some participants to evaluate the agents as cold and unfriendly which made these participants avoid situations which would result in an interaction: "*Cold and unfriendly and not liking any contact or maybe I wasn't making contact with them, I deliberately avoid to make a contact with them.* (P12111)". The agents were programmed to avoid collision with the participant, however, one participant viewed this action as a sign of unfriendliness when trying to interact with the agent: "*When I walked towards them, they turned and so in a sense those aren't the kind of a friendly reaction* (P41111)".

C.5.3 Participant responses and perceived agent attributes

Most participants recognised the agents as unexpressive or neutral however some participants interpreted the simple behaviours modelled in the agents to signify different states. In addition to being described as cold, unfriendly and indifferent, the agents were also described as annoyed, arrogant, grumpy, indifferent, preoccupied, purposeful and grumpy.

One of the participants perceived different emotional states in different agents in the virtual street: "*The Saudis were walking, just going about their business quiet calmly. And then the lady I felt was a bit annoyed when I bumped into her* (P21111)". The same participant went on to report how he identified with the agents when the experiment was coming to an end: "*It was funny, in the end when the light went down, and I had this fear of - I had to get out. Trying to find the way and I was lost. And then I saw a woman go by, and I wanted to stop her but I thought she was running as well and I had this impression/feeling that everybody was doing the same as I was* (P21111)". Another participant explained their behaviour in terms of the agent's perceived state (grumpy): "*The agents were just sort of kind of grumpy, I guess that is why I moved out of their way instead of seeing if they moved out of my way* (P51221)". Similarly another participant avoided interacting with the agents because in addition to perceiving the agents as cold and unfriendly, the participant did not identify common interests with the agents: "*I presumed these people were going to shops and buying stuff they were interested in and I am not very much interested in this stuff, so I wouldn't have great conversation with them* (P12111)". Some participants perceived the agents as being preoccupied: "*They just seemed very... in their own little world and they didn't care about what else was happening, they just seemed like they were on a mission there* (P12221)" or indifferent: "*They seemed to be indifferent to our presence* (P52121)".

More interestingly participants reported feeling realistic responses to interactions with the agents in the virtual street. For instance one participant (P21111) reported an urge to ask agents for directions when he felt lost and identified with the agents while another participant (P32211) reported that she did not interact with the agents for fear of a negative reaction. Some participants felt hesitant to break social norms and bump into the agents. For instance one participant did not carry out their plan of trying to test the agent: "*I thought of bumping into one of them there. I was not sure if that's a good idea* (P42221)". Another participant reported an incident in the virtual street with an oncoming agent: "*The aim was to to see if she was going to walk into me or walk around me. But then, when I was doing it, I kind of hesitated last minute. It is really anxious to stand still, for some reason, just kept swinging, left and right* (P11221)". In some cases, participants made an effort to avoid colliding with agents: "*From my experience before everybody I tried to bump into, everyone I tried to have contact with, they all moved to the side. I tried to do the same thing with her but I kind of stopped and it seemed like she did come at me. And then I jumped back and I was shocked that it had happened. And then again I laughed because it was a bit silly being scared of something that doesn't exist* (P31211)". A small number of participants even expected the agents to apologise either vocally or using gestures in a situation that can be described as virtual pedestrian rage: "*I remember bumping into someone. And I got a bit annoyed because he didn't turn around and say sorry* (P51121)".

Appendix D

Experiment on posture: Pictures

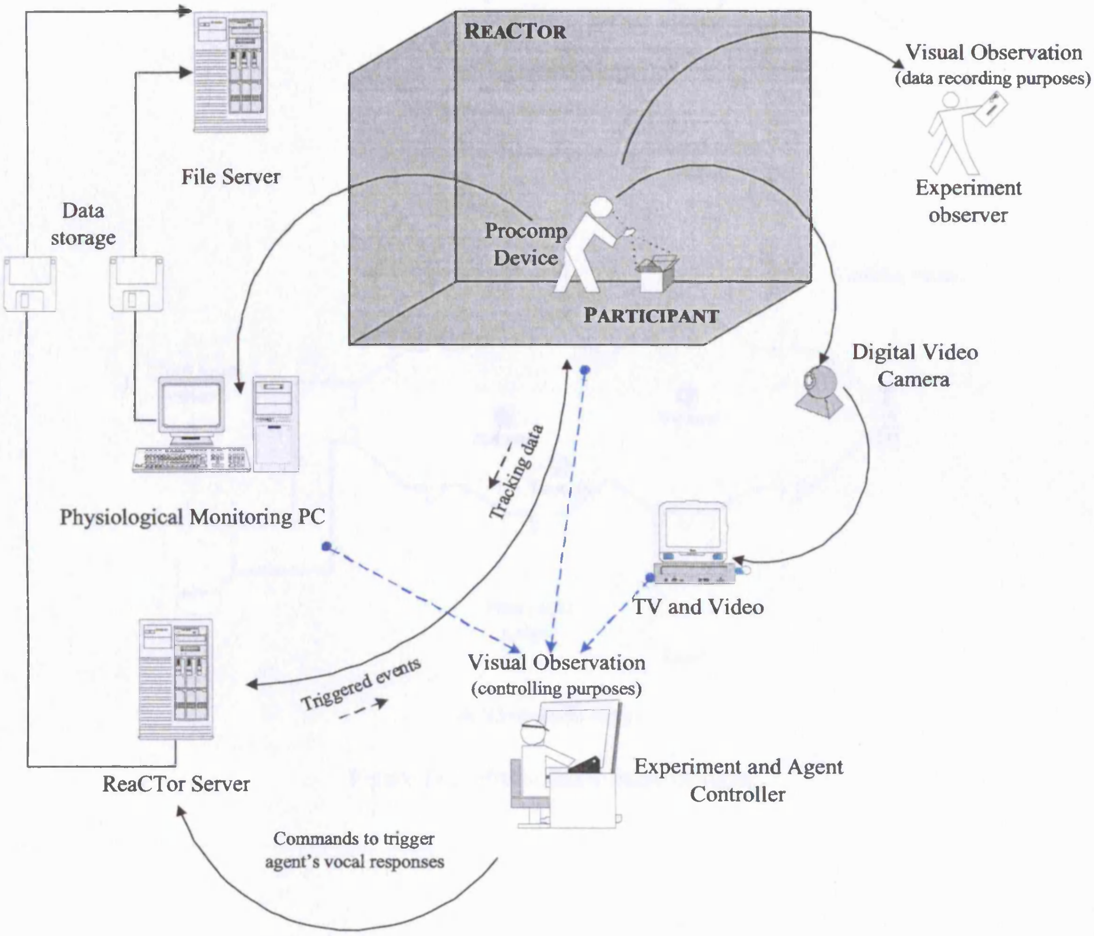
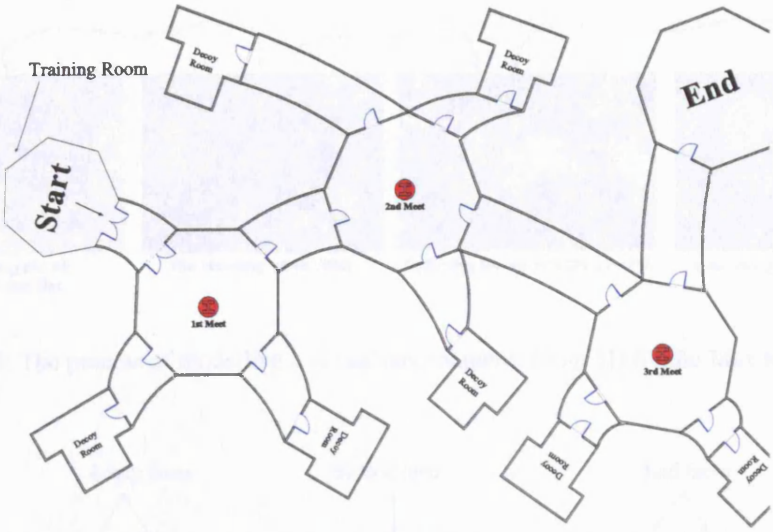
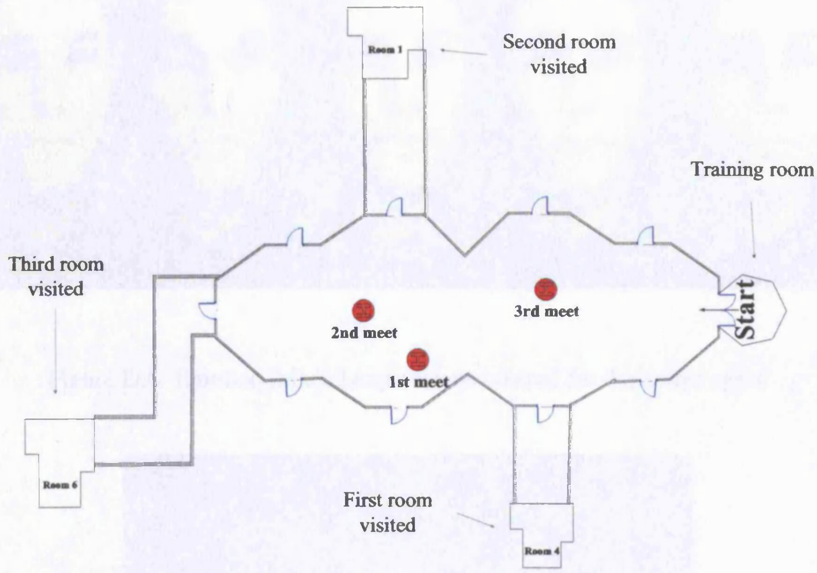


Figure D.1: Illustration of the setup during an experiment.



(a) Sequential design



(b) Centralised design

Figure D.2: Preliminary maze designs

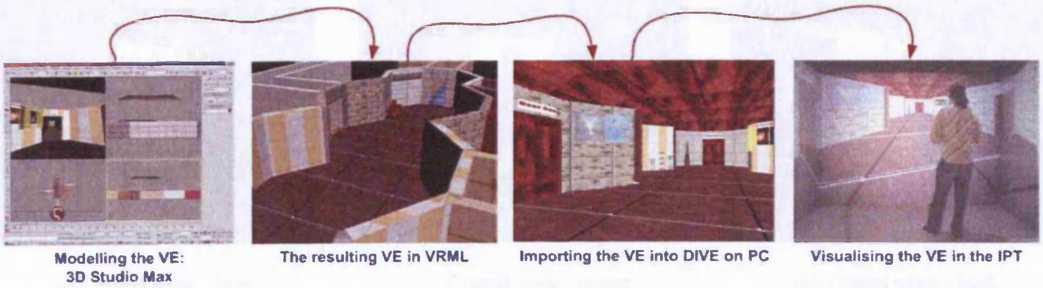


Figure D.3: The process of modelling a virtual environment: From 3D Studio Max to the ReaCTor

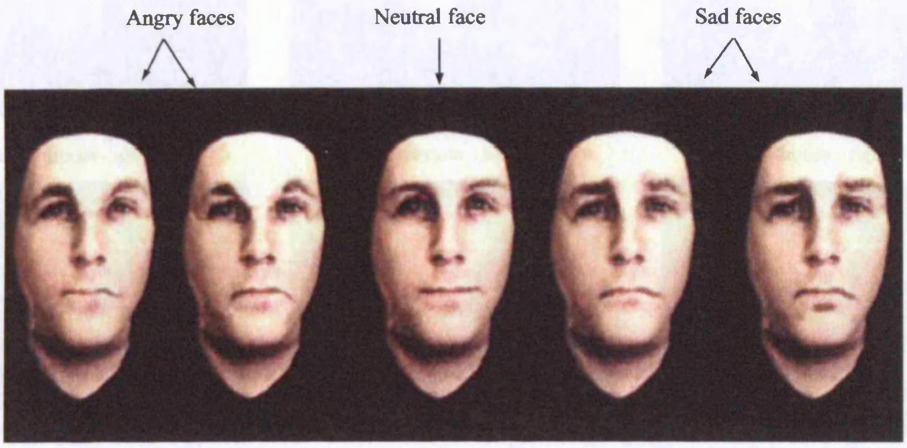


Figure D.4: Emotional facial expressions created for the active agent

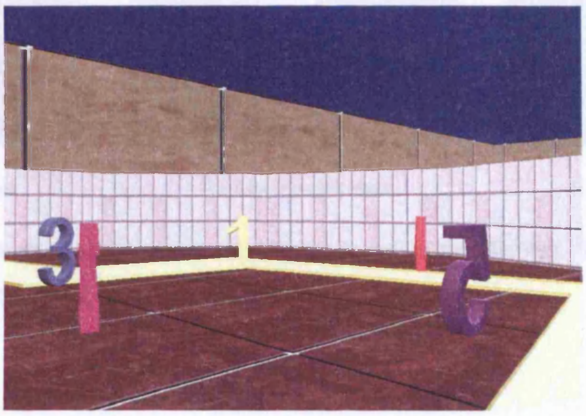
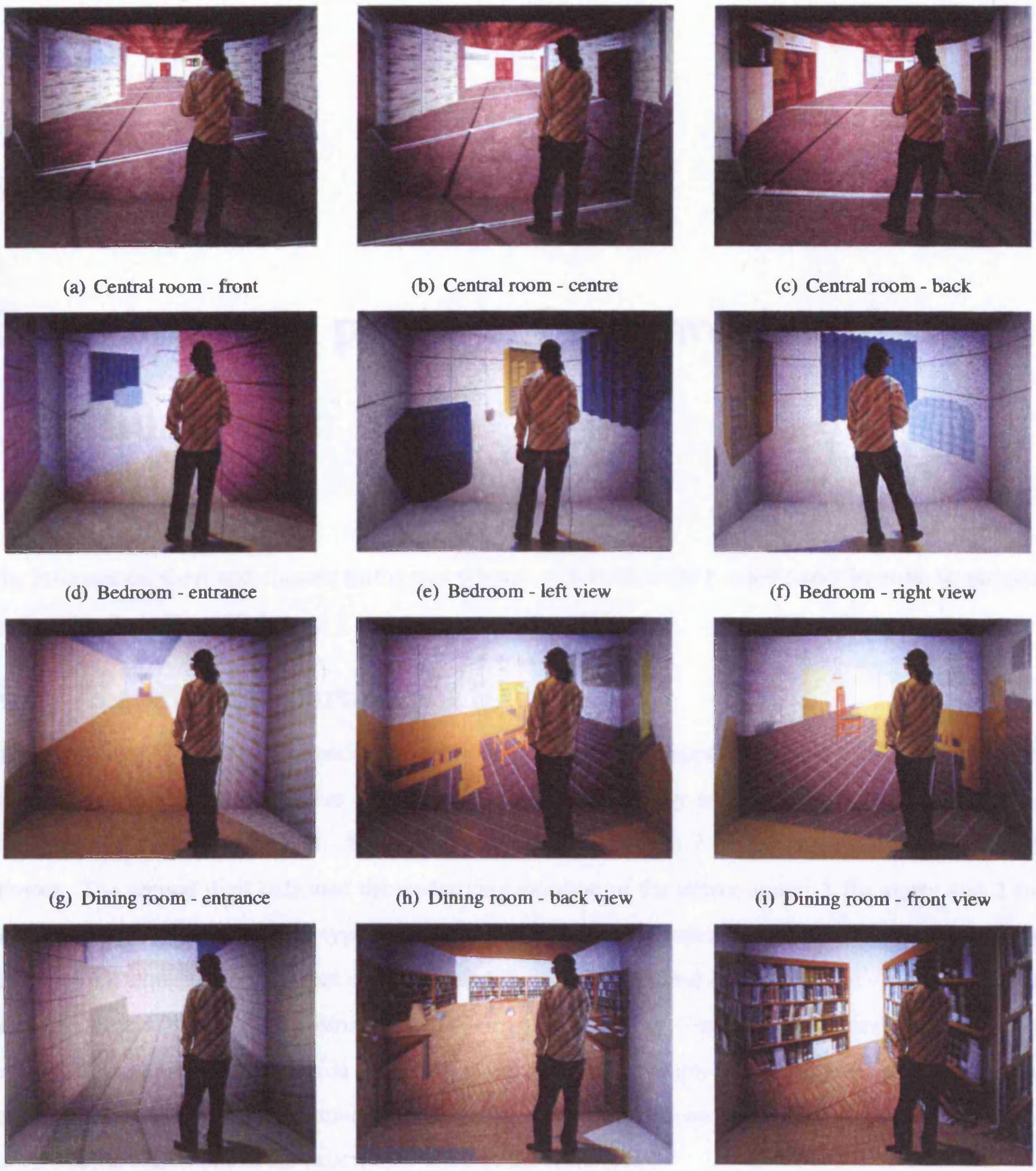


Figure D.5: The training environment with 3D numbers



(a) Central room - front (b) Central room - centre (c) Central room - back

(d) Bedroom - entrance (e) Bedroom - left view (f) Bedroom - right view

(g) Dining room - entrance (h) Dining room - back view (i) Dining room - front view

(j) Library - entrance (k) Library - left view (l) Library - right view

Figure D.6: More snapshots of a participant visiting the central room (D.6(a) - D.6(c)), bedroom (D.6(d) - D.6(f)), dining room(D.6(g) - D.6(i)) and library (D.6(j) - D.6(l)).

Appendix E

Experiment on posture: Paperwork and Questionnaires

The information sheet and consent forms was printed on formal letter-headed paper in order to provide contact details to the participants.

E.1 Decoding the participant ID

The 49 sessions (7 batches x 7 conditions) were randomised and mapped to a time slot in the schedule. Each session (and participant) was coded with a four-digit identity code based on the batch number and condition of the experiment. The first digit ranged from 1 to 7 in correspondence to the batch number. The second digit indicated the underlying emotion of the active agent: 1 for angry and 2 for sad. The third digit indicated the type of facial expressions - 1 for emotional facial expression and 2 for neutral facial expressions. The last digit indicated the type of postural cues displayed - 1 for emotional posture cues and 2 for neutral posture cues. For instance an identity code of 5121 translated to the fifth participant in condition 3, in which the active agent portrays an angry emotional state through neutral facial expressions and angry postural cues. Since the neutral conditions 4 and 8 were the same, the code for all neutral conditions in the experiment were of the form N122.

E.2 Information sheet for Participants

Thank you for participating in our study. This is one of a long series of studies into understanding the responses of people within virtual environments. This study has been approved by *University College London's Committee on the Ethics of Non-NHS Human Research.* Please read through this information sheet and feel free to ask any questions. The experimenters will answer any general questions, however, the specific aspects regarding this study cannot be discussed with you until the end of the session. The whole study will take about *an hour and a half.*

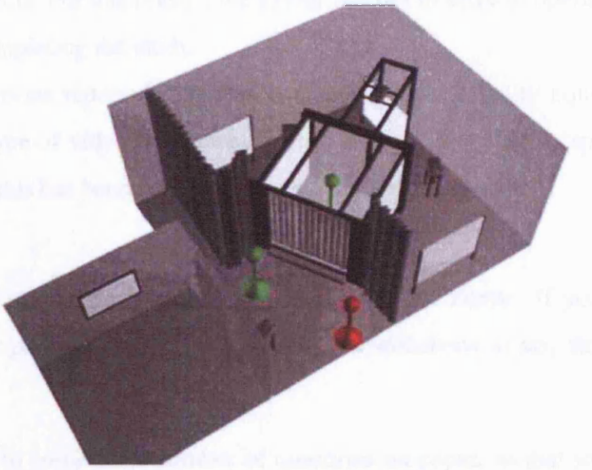


Figure E.1: Model of the VR system

You will be using the CAVE ^{TM}-like system called the ReaCTor. See figure above. The ReaCTor is a VR system made up of 3 walls measuring roughly 3m x 3m x 3m. You will wear VR glasses and be handling a tracking device similar to a joystick. The virtual reality viewing equipment can be worn over eyeglasses. You may be asked to take off your shoes in order to protect the virtual reality equipment. In addition to the tracking equipment used to navigate the system, you will also be fitted with physiological equipment designed to measure your heart rate, respiration and galvanic skin responses whilst you are in the immersive virtual environment.

In this particular study you will be asked to enter a building (maze) consisting of a central room with doors that leads to a number of surrounding rooms. You will be able to explore three rooms within the maze, however, which three rooms you explore will be decided by the people in the central room. You will be able to talk to the people in the building in order to get information about the door that leads to the room you should be visiting next. The people will always stay behind in the central room while you explore the three rooms. After you finish exploring each room, come back to the central room so that the people can tell you which door leads to the next room. Your task is simply to explore the rooms and remember the number (if any) of trash bins you see.

Information that we collect will never be reported in a way that specific individuals can be identified. Information will be reported in a statistical and aggregated manner, and any verbal comments that you make, if written about in subsequent papers, will be presented anonymously.

IMPORTANT

When people use virtual reality systems, some people sometimes experience some degree of nausea. If at any time you wish to stop taking part in the study due to this or any other reason, please just say so and we will stop.

There has been some research, which suggests that people using head-mounted displays might experience some disturbances in vision afterwards. No long-term studies are known to us, but short term studies carried out suggest that even after about 30 minutes of leaving a VR system, some people experience aftermath effects. For this reason, we advise you not to drive or operate heavy machinery for at least three hours of completing the study.

There have been various reported side effects of using virtual reality equipment, such as 'flashbacks'. Also with any type of video equipment there is a possibility that an epileptic episode may be generated. For instance, this has been reported in computer video games.

PROCEDURES

- You will be asked to read, understand and sign a **Consent Form**. If you sign it the study will continue with your participation. **Note that you can withdraw at any time without giving any reasons.**
- You will be asked to complete a number of questions on paper, so that we can try to understand your responses during the study.
- You will be fitted with sensors to measure your heart rate, respiration and galvanic skin responses.
- You will be asked to remove your shoes and switch off your mobile phone(s) before using the VR system.
- You will have a brief practice period with the help of the experimenters so that you can get used to the immersive virtual environment and learn how to navigate in it. You will then go into the maze as mentioned during which you will be videotaped.
- After the visit to the maze you will complete some more questionnaires about your experience.
- Finally there will be a short discussion with the experimenters about your experiences. The interview might be audio or video taped.
- **Thank you** for your participation. Please do not discuss this study with others for about three months, since the study is ongoing.
- Any other questions? *Please **ask** any questions* that come to mind at this point. After this read and sign the **Consent Form**.

In case you have any enquiries regarding this study in the future, please contact: Vinoba Vinayagamoorthy, Department of Computer Science, University College London, Gower Street, London WC1E 6BT +44 (0) 20 7679 7215 *Tel*, +44 (0) 20 7387 1397 *Fax*, V.Vinayagamoorthy@cs.ucl.ac.uk, http://www.cs.ucl.ac.uk/staff/V.Vinayagamoorthy

E.3 Consent Form

University College London

Department of Computer Science
Mel Slater
Professor of Virtual Environments

ID ______________________________

PROJECT	**EQUATOR**
Investigators	Vinoba Vinayagamoorthy, Mel Slater and Anthony Steed
Experiments	Ashwin Beeharee, Andrea Brogni, Richard Milton and Vinoba Vinayagamoorthy

***To be completed by volunteers*:**
We would like you to read the following questions carefully and **circle** your answers.

Have you read the information sheet about this study?	YES/NO
Have you had an opportunity to ask questions and discuss this study?	YES/NO
Have you received satisfactory answers to all your questions?	YES/NO
Have you received enough information about this study?	YES/NO
Which investigator have you spoken to about this study?	…………
Do you understand that you are free to withdraw from this study?	
• At any time	YES/NO
• Without giving a reason for withdrawing	YES/NO
Do you understand and accept the risks associated with the use of virtual reality equipment?	YES/NO
Do you agree to take part in this study?	YES/NO
Do you agree to be video taped?	YES/NO
Do you agree to be audio taped?	YES/NO
Do you agree to be physiological monitored?	YES/NO

I certify that I do **not** have epilepsy.

I certify that I will not be driving a car, motorcycle, bicycle, or use other types of complex machinery that could be a danger to myself or others, within 3 hours after the termination of the study.

Signed...**Date**.....................................

Name in block letters..

Investigator...

In case you have any enquiries regarding this study in the future, please contact:

Vinoba Vinayagamoorthy

Tel
Fax

Department of Computer Science
University College London
Gower Street
London WC1E 6BT

http://www.cs.ucl.ac.uk/staff/V.Vinayagamoorthy

Information that we collect will never be reported in a way that individuals can be identified. Information will be reported in aggregate, and any verbal comments that you make, if written about in subsequent papers, will be presented anonymously.

E.4 Personal information

ID ______________________________

Your Age	
How fluent is your English?	Basic ☐ Proficient ☐ Fluent ☐
Occupational status *(If other, please specify and also your area of interest)*	☐ Undergraduate Student ☐ Masters Student ☐ PhD Student ☐ Research Assistant/Fellow ☐ Staff - systems, technical ☐ Faculty ☐ Administrative Staff ☐ Other
Are you taking any medication? *If yes, please specify*	No ☐ Yes ☐
Did you consume more than 2 units of alcohol within the last 6 hours? *(2 units of alcohol = 1 pint of beer or 2 glasses of wine)*	No ☐ Yes ☐
Please state your level of computer literacy on a scale of (1...7) (novice) 1 2 3 4 5 6 7 (expert)	
Please rate your level of experience with computer *programming* (novice) 1 2 3 4 5 6 7 (expert)	
Have you ever experienced 'virtual reality' before? (no experience) 1 2 3 4 5 6 7 (extensive experience)	
How many times did you play video games (at home, work, school, or arcades) in the last year?	Never ☐ 1 - 5 ☐ 6 - 10 ☐ 11 - 15 ☐ 16 - 20 ☐ 21 - 25 ☐ > 25 ☐
How many ***hours per week*** **do you spend playing video games?**	0 ☐ < 1 ☐ 1 - 3 ☐ 3 - 5 ☐ 5 - 7 ☐ 7 - 9 ☐ > 9 ☐

E.5 Agent emotion assessment questionnaire

The first item on the questionnaire was a descriptive style question which would be covered later in the semi-structured interview. This was followed by a set of three pairs of questions which were included partially to check if the agents responded correctly towards the participants.

Please note down your impression of what you thought the people in the maze were doing?		
Item	**Person on YOUR left**	**Person on YOUR right**
Short description		
He responded to me	Yes/No	Yes/No
I interacted with him	Yes/No	Yes/No

Table E.1: Section 1 of the agent emotion assessment questionnaire

The second section in the questionnaire dealt with the participants perception of the agent's emotional state towards each other during the approach period of the participant-agent interplay.

Choose ONE of the following statements to describe the people in the maze towards EACH OTHER		
Tick one of the following statements that is most applicable to each person in the maze	**Person on YOUR left**	**Person on YOUR right**
He seemed surprised at the other person		
He seemed afraid of the other person		
He seemed angry at the other person		
He seemed happy at the other person		
He seemed disgusted at the other person		
He seemed sad at the other person		
He seemed neutral towards the other person		

Table E.2: Section 2 of the agent emotion assessment questionnaire

The final section in the questionnaire dealt with the participants perception of the agent's emotional state towards the participant during the interaction period of the participant-agent interplay.

Choose ONE of the following statements to describe the people in the maze towards YOU		
Tick one of the following statements that is most applicable to each person in the maze	**Person on YOUR left**	**Person on YOUR right**
He seemed surprised at me		
He seemed afraid of me		
He seemed angry at me		
He seemed happy at me		
He seemed disgusted at me		
He seemed sad at me		
He seemed neutral towards me		

Table E.3: Section 3 of the agent emotion assessment questionnaire

E.6 The SUS questionnaire

SUS Presence:(n=5) Response variables were collected using the questions adapted from Slater and Steed (2000);

1. Please rate your sense of being in the maze, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place.
2. To what extent were there times during the experience when the maze was the reality for you?
3. When you think back about your experience, do you think of the maze more as images that you saw, or more as somewhere that you visited?
4. During the time of the experience, which was strongest on the whole, your sense of being in the maze, or of being in the real world of the laboratory?
5. During the time of the experience, did you often think to yourself that you were just standing in a laboratory or did the maze overwhelm you?

E.7 The new Presence and Copresence questionnaire

The presence/copresence questionnaire based on the concept of operational presence as defined by Sanchez-Vives and Slater (2005).

New Presence:(n=6) The first part contains the following questions, each on a 1 to 7 scale where '7' indicates 'higher presence';

1. How much did you behave within the maze as if the situation were real?
2. How often did you find yourself automatically behaving within the maze as if it were a real place?
3. How much was your emotional response in the maze the same as if it had been real?
4. How much were the thoughts you had within the maze the same as if the maze had been a real situation?
5. How much were you thinking things like 'I know this isn't real' but then surprisingly finding yourself behaving as if it was real?
6. To what extent were your physical responses within the maze (e.g., heart rate, blushing, sweating, etc.) the same as if the maze had been a real situation?

New Copresence:(n=6) The second set of questions concerned the response to the agents and were also rated on a scale of 1 to 7;

1. How much did you behave as if the characters were real?
2. How much did you find yourself automatically behaving as if the characters were real?
3. How much was your emotional response to the characters as if they were real?
4. How much were your thoughts in relation to the characters as if they were real?
5. How much did you have physical responses (such as change in heart rate, blushing, sweating, etc.) to the characters as if they were real?
6. How much were you thinking things like 'I know this person isn't real' but then surprisingly finding yourself behaving as if the character was real?

E.8 Modified version of the SAD questionnaire

1. I felt relaxed
2. I wanted to avoid the situation
3. It was not easy for me to relax with those strangers
4. I had no particular desire to avoid them
5. I found it upsetting
6. I felt calm and comfortable
7. I wanted to try to avoid talking to them because I didn't know them
8. I took the chance to meet new people
9. I felt nervous or tense
10. I was nervous with people because I didn't know them
11. I felt relaxed even though I was with those people
12. I often wanted to get away from those people
13. I felt uncomfortable because I didn't know them
14. I felt relaxed even though meeting these people for the first time
15. Having to introduce myself made me tense and nervous
16. Even though it was full of strangers I felt comfortable to go in anyway
17. I wanted to avoid walking up to and joining those people
18. I talked willingly with them
19. I often felt on the edge when I thought about talking to them
20. I tended to withdraw from the people
21. I didn't mind talking to them
22. I was seldom at ease with them
23. I was trying to think up excuses to avoid being with them
24. I tried to avoid this situation
25. I had no problem to go over to them
26. I felt it easy to relax with the other people

E.9 Summary of the interview questions

1. What was your impression of the people when you first saw them?
2. What did they seem to be doing?
3. What were they talking about?
4. How did this change over the three consecutive meetings?
5. Overall, how would you describe the behaviour of the people towards you?
6. How real did the people seem to you?
7. How real was their behaviour to you?
8. Did they know each other?
9. Did you feel you were with other people?
10. How did they react to you?
11. Can you describe their behaviour when they saw you?
12. Were they aware of you? How do you know that?
13. How would you rate your response towards the people?
14. How did it affect you when a person looked at you?
15. On the whole, how did you respond to the virtual people?

E.10 Interview Structure

Figure E.2: Agents used in the maze

The planned structure of the interviews is outlined below.

(GENERAL):

- What did you think was going on in the maze?
  - What made you think that?
  - What specifically?
- How would you describe what went on in the maze to a friend?
  - How specifically?
  - What makes you think that?
  - Let me see if I understand you. You are saying that...

(FEELING OF PRESENCE):

- To what extent was your experience that of being in a maze rather than in a laboratory?
  - What about the maze made you think it was similar? Is there anything specific?
  - Alt: What about the maze made you think it wasn't similar?
- How did this change over time? (Figure E.3)
  - You are saying that...
- To what extent was your behaviour/response appropriate to the situation portrayed in the maze (*rather than the behaviour of being in a laboratory study*)?
  - How would you have behaved in a maze? What was similar in the maze?
- How did this change over time? (Figure E.3)
  - You are saying that...

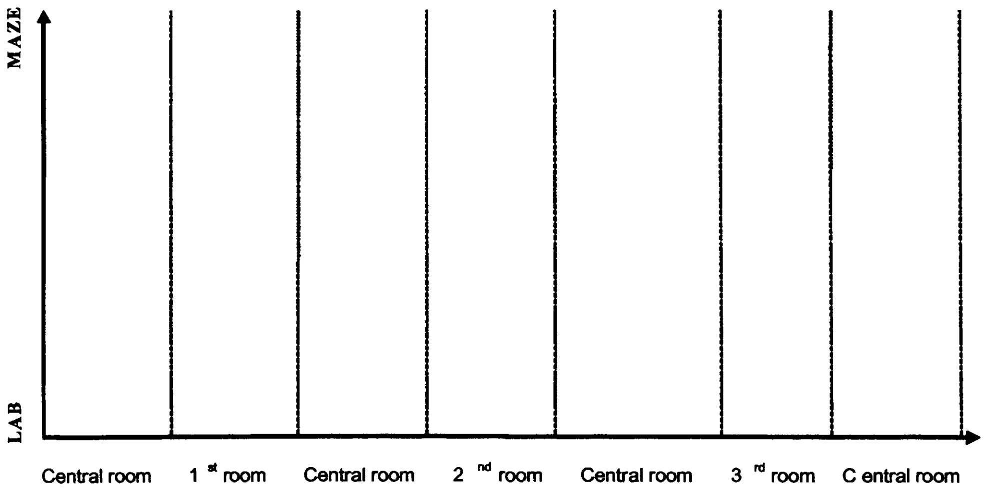


Figure E.3: Qualitative visualisation template used in interviews

(FEELING OF TRANSITIONS TO REAL):

- Did you ever experience a change so that although your experience had been in the maze, you were suddenly back in the laboratory?

  - What specifically made you experience this change?

- How quickly did you recover (felt being in the maze again)?

  - Were there different intensities of change and different amounts of time it took you to recover?

- Is there anything more that you can say about this experience?

  - You are saying that...

- Give another example of when this happened?

(INTERACTION WITH PEOPLE):

- What was your impression of the people when you first saw them? What did they seem to be doing? What were they talking about? How did this change over the three consecutive meetings?

  - What makes you think? How do you know that?

  - Let me see if I understand you. You are saying that...

- Overall, how would you describe the behaviour of the people towards you?

  - What makes you think? How do you know that?

  - Let me see if I understand you. You are saying that...

- Overall, how did this make you feel?

  - What specifically makes you think this? How do you know?

- How real did the people seem to you? How real was their behaviour to you?

  - What specifically about their behaviour makes you say that?

  - Let me see if I understand you. You are saying that...

- Did they know each other? Did you feel you were with other people? In their company?

  - What specifically makes you think that?

  - How do you know?

- Did they react to you? How? Can you describe their behaviour?

  - How do you know that?

- Were they aware of you? How? Can you describe their behaviour?

  - How do you know that?

- Were there any one particular person to whom you had a different response? Stronger response? Special response? ....

  - Which ones specifically? What about the person made you think that?

- How did it affect you when a person looked at you? Smiled at you? Talked to you? What was your response? Talking to you? Looking at you? On the whole, how did you respond to the virtual people?

  - Is that how you would respond/behave in a real maze?

- To what extent are you surprised by your own reactions:

  - To being in the maze?

  - To being with the people?

  - Let me see if I understand you. You are saying that...

(FURTHER INFORMATION):

- Is there anything else you would like to say about:

  - Your overall experience in the maze

  - Your reactions to the characters

  - Your thoughts and feelings during the experience in the maze

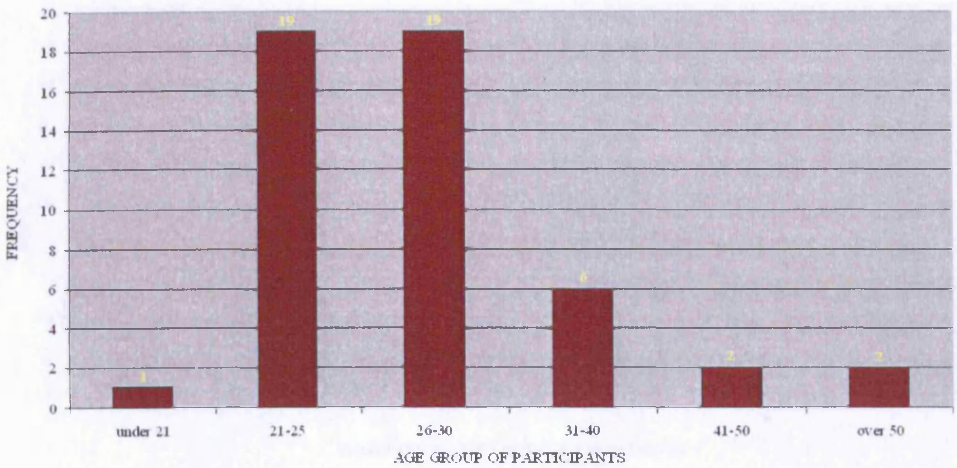
  - Your thoughts and feelings after you left the maze

  - Anything else...

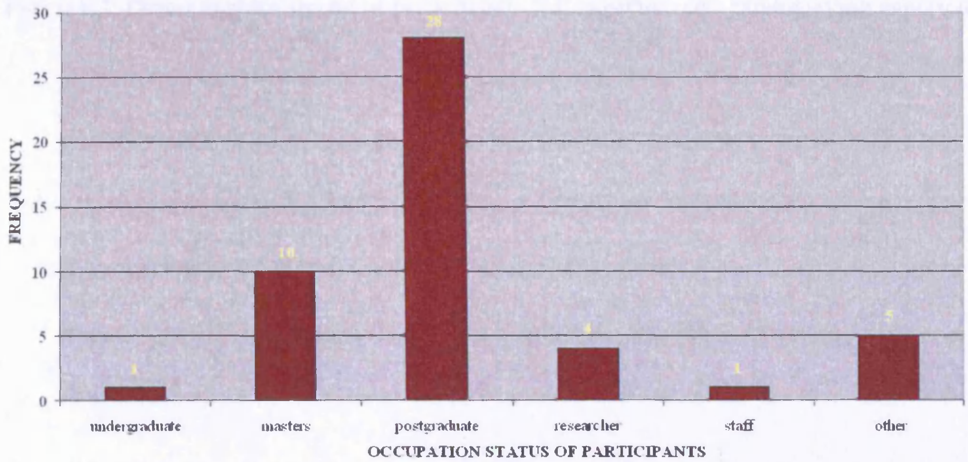
Appendix F

Experiment on posture: Analysis and Results

F.1 Population

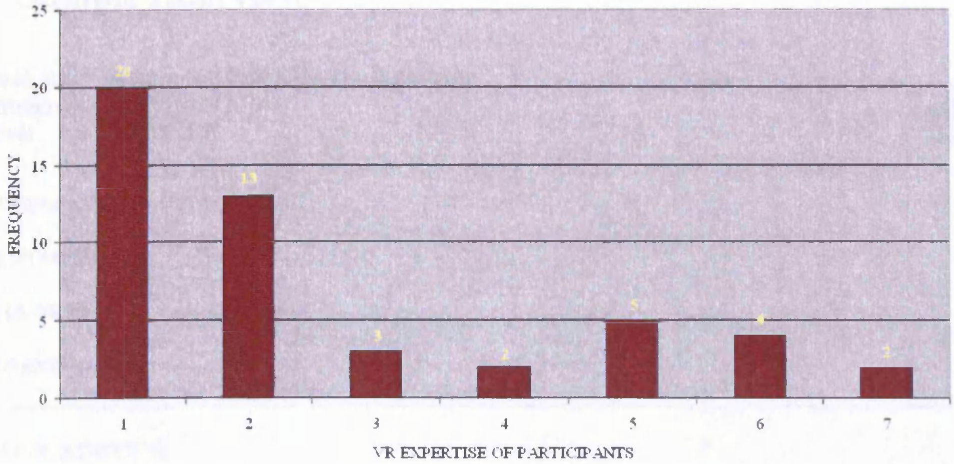


(a) Age Group

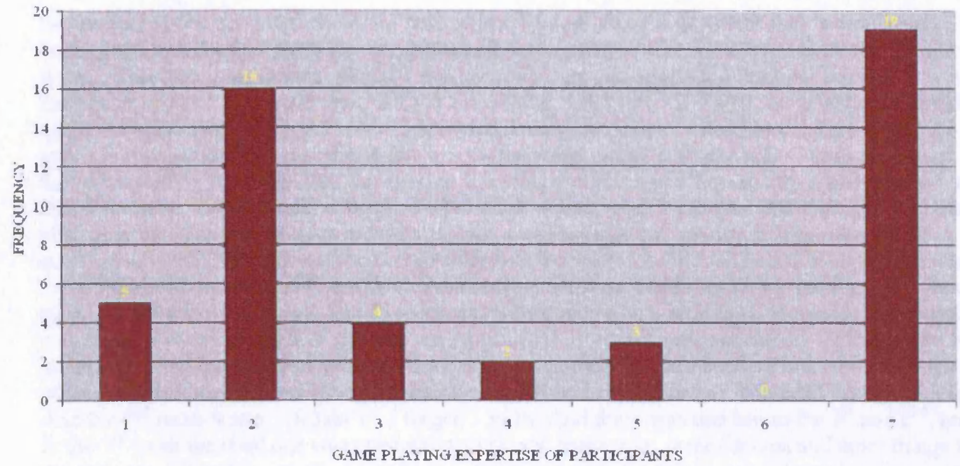


(b) Occupation status

Figure F.1: Demographics spread of participants: Age and Occupation



(a) Experience in VR systems



(b) Game playing experience

Figure F.2: Demographics spread of participants: VR expertise and game playing experience

F.2 Sample Interview

Experiment: Posture and Emotion: The maze 2005
Environment: The Virtual Maze
Project: EQUATOR

PARTICIPANT ID	[4][1][1][1]
TAPE NUMBER	[07]
SIDE (A OR B)	[A]
COMMENTS	[....]

[START OF INTERVIEW]

E Participant id 4111. So if you start from the top, what did you do in the maze. Describe what you did. Did the doors open

P Doors opened and I went to the main hall, where I found these 2 guys and they were....yeah...in the beginning I thought they were like computer software, I mean...then I realise, I think that they actually were....were you commanding them? I don't know, I am not sure but I think so because it would be really complicated to program and answer everything. Like in ....if I said, for example, I could have made the questions really different from what I did...so it was recognised...so...then it is really interesting because they...questionnaire was really based about with my relationship with them and I thought that they were just like direction to knowing the place and not exactly some people I should meant to know and to speak with. And after I asked him, cos' there was just the right one that was talking to me. The left one never talked to me. Cos' he was not exactly neutral like I placed in the questionnaire but there was not the correct adjective, he was more like different, so more like whatever I ask, he would answer and [inaudible]. And he was showing the direction in which room I should go after, I went to these 3 rooms. The 1st room was a bedroom with a little fan and outside was a garden. It was funny because I think I crossed the wall. And I wasn't suppose to do it. But yeah, it was interesting. And then I came back and I went to the guy again and he asked me, told me to go to the other room that was, I think. The 2nd room was, what was the 2nd room? The third room was a library. And the 2nd room was a....I didn't...I forget....yeah. And there was one bin in the 1st and 2nd, and two in the 3rd. yeah the third one was most interesting cos' there was...more details and more things to see, the only one I actually....really needed to go down and look for things, see if I was missing some bins and behind....

E did you find any?

P Yeah...it was really cool. Last one was the most interesting and yeah, as I in the questionnaire. I think the only problems were like, apart from crossing the walls which was a bit weird, I felt strange about crossing the walls, crossing the objects, crossing the people as well. Other thing was sometime, like some slow motion like whenever you....doesn't make me feel sick exactly but makes me feel uncomfortable like....

E ok, was this constant throughout the experiment or?

P A little bit especially when there was more detail. The [inaudible] been like too loud or something. It was less...the library was the one that I felt most [inaudible] but apart from that, it was a really interesting experience.

E So in the whole of your experience in there, were you more aware of....rather were you more in the maze or in the laboratory

P more in the maze, I believe

E ok, did this change over time?

P no, I just saw the laboratory in the end and I turned a little bit to the left more than I should probably and then I saw the....yeah, apart from that, I felt I was in the maze.

E ok

P all the time

E if I were to ask you to relate this to a qualitative graph. So this would be your sense of being in the maze and this would be your sense of being more in the laboratory. This is your time scale, this is your first visit to the central room, then 1st room, central room, 2nd room... which was the dinning, central room, third room was the library, if you could plot your sense of being. So were there some places where you felt you were more in the maze, did your sense of being in the maze increase or decrease over time? Did something happen in the middle or in the end? That made you feel more in the laboratory than in the maze?

P I think.... no.... its easier for me because I was always like this kind of electronic stuff and digital things in ... so I always like computer. So everything I felt like in the maze. I got [inaudible] probably just in the end, when I turned in that room, it would be like.... how would can I do it?

E So you can either draw a line or you can draw points and join the points, which I can do after all the...

P its like this

E ok ok... so you were pretty high when you went into the central room the first time, ok... the doors opened, you were completely in the maze

P just... yeah I... I think... maybe it should be like this, started to get down here because in the... as in the third room, there were more details but even with more details, the things like the books are not, the papers, they were not enough details on them so sometimes I felt like I was looking and I was trying to realise what was that?

E So you were trying to read the book title?

P yeah yeah... it was like. There was one that was looking like harry potter. It was looking at it.... there was a little blur but ... yes

E ok so in terms of your behaviour and responses, displayed in the maze. Say I want to... suppose I build the maze in a physical sense, a real maze with real people, gave you the same task under the same experimental situation. How do you think, how different or how similar do you think your behaviour would have been in terms of the way you approach the people. Exploring the place... observing the place

P you mean, if they were real person?

E yes

P Ah, it would be completely different.

E ok so if it was going to be different, it would be more towards the lab and if it was similar to what you would normally do, then it would be near the maze. So same kind of thing, so in the central room, you would wonder, you would think about things like how you approached the people, how you interacted with the people, how you looked around in the central room. Where in the first room, you would think about just the exploring and looking for trash bins and judge whether your behaviour in the virtual space here was similar or different to how your behaviour would have been, had it been a real situation?

P Do you mean, I think I didn't understand very well. Do you mean, if there were not computer based and they were real people, I think it would be more realistic probably.

E No no, I mean if I were to take this virtual place and build it somewhere in the real, so you had a real maze, you had real people standing in there, but you were given the same task to go into this building and find the, so I want you to think about how similar or different your behaviour would be in comparison to what happened here...

P ok now I understand....

E so let us....from the beginning, let us think about for example, the doors opened, did you look around in the central room? Or did you immediately go to the people? What did you do here?

P no I looked around....

E ok so you looked... is that something you would do normally if it was a real place?

P because I don't know the place, so

E looked around

P look around

E ok so in that case I would say that well, its similar to what you did, so therefore its high. Right? Now once you started going towards the people, how did you behave?

P Yeah I think it changed a little bit because you have no facial expression in the....

E No I am not talking about their behaviour....Their behaviour, we will talk about later. I am talking about how you behaved?

P How I behaved?

E Yes how you went up to them. Did you saw hello first or did they say hello first? How did it happen?

P I felt it was like... oh....they said something before

E ok is that what you think will normally happen if you were in a real situation? Do you think...

P Yeah probably because when you get to a different place, where you don't know the place, you don't know the people, normally people will say something.

E so after that, how do you think the interaction went? Was it how you would normally think of an interaction or was it very different to what you think your interaction with the characters would have been?

P Think it would have been very different. What I felt it was like, I felt they were really arguing with each other. The right one like shouting at the other one. And then I got there, yeah it was interesting, I thought it was pretty real and I felt like I was interrupting something. And then they... yeah it was really cool but from what having their facial expression, the hearing them arguing and when I was getting to them, it was cool to....and I felt like embarrassing maybe. Should I be here.

E yes but how about in terms of how you talked to them? Do you think that something... is that how you would talk to real people, the way you talked to...

P yes yes but I must be objective because I was afraid about them cos I didn't know whether it was computer software or if there were people talking for them. So ...I was concerned about that particular thing, I didn't know if they will answer to anything that I should ask.

E So in terms of... if they had been real people, would you have talked about something differently? Would you have approached...

P sure yeah

E what would you have done differently?

P I should, probably I ask about the place, ask about that.

E So you feel you would have had more chit chat?

P Yes yes

E ok so in that case, well your interaction was very different, and therefore it was somewhere down here. Ok so this is what I want you to think about when you are doing this graph. So start up

P yeah ok

E so you went into the central room, you looked around, that was similar to what you would normally do, so therefore it is high. But you feel your interaction was not as what you would do normally, therefore it was low and then afterwards...

P so it would be... I think it would be... how do you mean? Dots like this...

E yes

P I think it will move like this again and like this...

E but in the first room, you are looking around? There is no interaction, this is when you are in the 1st bedroom...

P ok here....

E so in the bedroom, would you have done anything differently?

P yeah I think no no no....so its like maze....so its like all the rooms would be more like maze and all central room would be more like cos' ....

E interaction was not as....

P yeah not as...

E in the central room, the final end, was there...

P no no...cos' there were nobody

E ok

P so it would be....

E these one on the tops...

P like this....

E ok and this would again go up and would just stay down?

P yeah go up....yeah

E was there anything, other than you walking through the walls, was there anything that made you suddenly became aware of the laboratory while you were in the maze?

P yes when I ...in the end where I....

E so you saw the....was that because you turned around and saw the empty... ok... anything else?

P no the lack of details in the objects

E that is in the library room

P yeah especially cos' the other ones, like they were not completely furnished and there were not too much detail to notice... so not much...

E how quickly did you feel you recovered from these things that made you remember the laboratory more? How quickly did you get back in the maze and continue doing whatever...

P immediately... immediately, I just like look and that ... trying to see, wow this looks like harry potter but it is not pretty real but then I turn around thinking other than getting you...

E were some of them worse than the others?

P no it was just cool, when I saw like more detail... .a lot of pieces of papers shredded... when I crouched, I saw like bag on the floor

E ok

P under the table, it was cool... it was really pretty real....

E ok so when you walked into the central room, what was our impression of what they were doing initially?

P When I walked into the central room....

E before you went up to them, what did you think they were doing? What did they look like they were doing?

P I don't know... talking... .they were talking but I don't know exactly because I didn't pay attention.

E ok did you know whether it was a positive or negative or neutral... what they were talking about?

P I think they were positive in the beginning and after that the right guy started to complain about the left guy because of something he had done.

E ok so... .you have said that was trying to instruct the other but the instruction... .instructor was not very patient... .? When you say instructor?

P the right one

E the person on the right... would that be the person with... .so...

P so yeah the....

E ok and did this change over the 3 consecutive meetings?

P no... it was... yeah yeah... I think some of the... .thinking sometimes the guy was more angry and sometimes he was more relaxed but definitely the left one was completely afraid of the right one...

E what makes you say that? What gives you the impression that one of them was?

P he was ... the position ... .was all down and....

E who the person on the left or

P left...

E ok and what else gave you the impression of?

P the right guy shouting at the...

E ok why do you think he was shouting....was it just loud or

P yeah I think... yeah yeah

E and how did their behaviour towards you seem indifferent?

P indifferent... yes, the left one completely neutral because he didn't even ....I don't know, I think because of his position, he was wearing a suit and he was like just all the time like look like he was commanding the other one. Probably that is why I went to him

E ok did you deliberately made it a point to go to the person in the suit?

P Yeah I think so

E from the very start or?

P from the start

E from the start ok...

P probably yes... so I think it was neutral, the left one I think was neutral because I didn't reach him...

E Ok so what gave you the... so what happened when you went up to the people? How did they react to you? Did they react to you?

P Yeah... its difficult to you. Because they don't have expressions. You can not see exactly what people feel about so its little weird....

E reacted to you initially mean, when you walked up to them, what happened?

P yeah yeah they turned to me

E ok so both of them turned or just one?

P As far as I remember, just the right one, but I don't know maybe I think that I went straight to the right one, I don't remember the left...

E Did you try to talk to the left one at any point?

P no that must be the reason... now I am thinking about it, that may be....the left one was left behind....

E so you went to the right person and he reacted to you, he turned towards you? Once you started having a conversation with this person, did you feel that they were aware of where you were and of you being there?

P yes

E what gave you this impression

P they turned to me and talked with me

E and what else, did they smile at you or look at you or

P I don't think, at least they didn't notice... because I think they didn't. They have this possibility like

E which possibility?

P to smile and you know, I didn't feel….maybe I didn't notice but I don't think so

E ok how about looking at you?

P looking yes

E both of them or just the one?

P just the right one because the other one was ignored…

E ok, is that because you didn't pay attention to him or

P yes I didn't pay attention, I think

E when you were walking up to them, did you feel that they knew each other

P uh hmm

E and what gave you this impression?

P they were looking at each other and the right one was talking something

E and when you were with them, did you feel you were in the company of two people? Or the people?

P yes yea yeah

E ok and what specifically make you think that you were in the company of other people?

P because they were talking and they were moving like…

E and what gave you the impression that the person on the right was talking to you?

P impression…the sound

E ok how do you know he was…where the sound was coming from?

P that is interesting, maybe because of the way, I think he was not standing still, he was just moving his arms or I don't know if I am getting high here but I think it was that….

E ok how about….so did you have a different response to one of the people or did you have the same type of response to both?

P that's what I told you before…I think the reason the left one was neutral about me was because I didn't reach him

E so you had a stronger response to the person on the right?

P yeah

E ok and when you went…ok so you made it a point to go to all the 3 times to this person on the right?

P hmmm

E ok, have you been in VR before?

P no

E To what extent are you surprised by your reaction to the maze and the people?

P Yeah…I am …yeah I am surprised…because I didn't expect to feel a certain type I felt…to feel like I was really in the maze instead of in the lab and

E ok and you felt like you were….

P yes

E is there anything you want to add?

P what about the left guy….

E what you want to ask questions about the program…

P yeah

[END OF INTERVIEW]

F.3 Code Descriptions

Active	"Perceptions of the Active agent"
EInstructions	"Expected to get instructions from agents"
ENoComplexResponses	"Expected no or limited responses from the agents"
ENoQuery	"Expected no queries from agents"
FactorAffectBody	"Experiment Factor: Affective Postural Cues"
FactorAffectFace	"Experiment Factor: Affective Facial Cues"
FactorAngry	"Experiment Factor: Angry condition"
FactorNeutralBody	"Experiment Factor: Neutral Postural Cues"
FactorNeutralFace	"Experiment Factor: Neutral Facial Cues"
FactorSad	"Experiment Factor: Sad Condition"
PAcknowledgedMe	"Agent acknowledged me"
PActiveAgitated	"Active agent was agitated"
PActiveAngry2Participant	"Active agent was angry towards participant"
PActiveAnsweredMe	"Active agent answered me"
PActiveBodyLanguge2Other	"Active agent's body language to the other"
PActiveCalm	"Active agent was calm"
PActiveComplaining	"Active agent was complaining"
PActiveForceful	"Active agent was forceful"
PActiveForthcoming	"Active agent was forthcoming"
PActiveFriendly2Participant	"Active agent was friendly to the participant"
PActiveHigherAuthority	"Active agent was higher in authority"
PActiveImpatient	"Active agent was impatient"
PActiveInteractedMore	"Active agent interacted more"
PActiveLookedAtMe	"Active agent looked at me"
PActiveMoving	"Active agent was moving"
PActiveNotForthcoming	"Active agent was not forthcoming"
PActivePleasantPolite	"Active agent was pleasant and polite"
PActiveResponsive	"Active agent was responsive"
PActiveSad	"Active agent was sad"
PActiveSpoke	"Active agent spoke"
PActiveTellingoff	"Active agent was telling off passive agent"
PActiveTookCharge	"Active agent took charge"
PActiveUnimpressed	"Active agent was unimpressed"
PAgentBehaviour	"Perceived Behaviour of the agents"
PAppearanceUnreal	"Perceived appearance of agents was unreal"
PArgument	"Agents were having an argument"
Passive	"Perception of the passive agent"
PAttitude2Participant	"Perceived attitude towards the participant"
PAwareOfMe	"Agents were aware of me"
PBehaviourUnreal	"Agent behaviour was unreal"
PBodyLanguage	"Perceived body language of the agents"
PBodyLanguageUnReal	"Perceived body language was unreal"
PColdBehaviour	"Agent behaviour was cold"
PComputerGen	"Agents looked computer generated"
PDebating	"Agents were debating"
PDisagreement	"Agents were in a disagreement"
PDiscussion	"Agents were in a discussion"
PEngrossed	"Agents were engrossed"
Perception	"Participant's perception"
PExtremeBehaviour	"Agent's behaviour was extreme"
PFacialMovement	"Agent's perceived facial movement"
PFeltReal	"Agents felt real"
PFeltUnreal	"Agents felt unreal"
PFriendly2Participant	"Agents were friendly towards the participant"
PGreeting	"Agents greeted participant"
PImaginaryRoles	"Agents had imaginary roles"
PKnowledgeOfParticipant	"Agent's had knowledge of the participants"
PLimitedInteraction	"Had a limited conversation/interaction with the agents"
PLookedAtMe	"Agents looked at me"
PMisTimedResponse	"There was a delay in the agent's responses"
PMoving	"Agents were moving"
PNeutral2Participant	"Agents were neutral towards the participant"
PNoPoint	"There was no point in talking to the agents"
PNoticedMe	"Agents noticed me"
PPassiveAfraid	"Passive agent was afraid"
PPassiveAgitated	"Passive agent was agitated"
PPassiveDisturbed	"Passive agent was disturbed"
PPassiveNoAcknowledgement	"Passive agent gave no acknowledgement"

PPassiveRelaxed	"Passive agent was relaxed"
PPassiveShy	"Passive agent was shy"
PPassiveSilent	"Passive agent was silent"
PPassiveStanding	"Passive agent was standing still"
PPassiveSuprised	"Passive agent was surprised"
PPassiveTurnedLater	"Passive agent turned later"
PPassiveUninterested	"Passive agent was uninterested"
PPolite	"Agents were polite"
PPosition	"Spatial position in relation to the other agent"
PPredictableBehaviour	"Agent's behaviour was predictable"
PQuickResponse	"Agent's responses were quick"
PReactiveBehaviour	"Agent's behaviour was reactive"
Precognition	"Agents recognised participant"
PShocked	"Agents were shocked"
PSpokeFirst	"Agents spoke first"
PStoppedArgue	"Agents stopped arguing when participants walked there"
PStoppedTalk	"Agents stopped talking when participants walked there"
PSurprised	"Agents were surprised"
PTalking	"Agents were talking"
PTurnedToMe	"Agents turned to me"
PUnEasyInteraction	"Participant-agent interaction was uneasy"
PUnfriendly2Participant	"Agents were unfriendly to participant"
PUnhelpful	"Agents were unhelpful"
PUnInterested	"Agents were uninterested"
PUnsmoothAnimation	"Agent's behaviour animation was unsmooth"
PUnWelcoming	"Agents were unwelcoming"
PUpsetAtIntrusion	"Agents were upset at participant's intrusion"
PVocal	"Agent's vocal responses"
PWaitedMeSpeak	"Agents waited for me to speak"
RAmused	"Participant felt amused"
RDisappointed	"Participant felt disappointed"
Realistic Responses	"Participant's lifelike responses"
RFeltNervous	"Participant felt nervous"
RFeltNoAnxiousness	"Participant felt no anxiousness"
RIncreasedPresence	"Participant felt increase in presence"
RIntimidated	"Participant felt intimidated"
RIntrusion	"Participant felt intrusive"
RIrritated	"Participant felt irritated"
RKnewActiveWldSpeak	"Participant knew active agent would speak"
RKnewNotReal	"Participants knew that the agents were not real"
RNoTrust	"Felt no trust towards the agents"
RPaidNoAttn	"Paid no attention to the agents"
RPaidNoAttn2Passive	"Paid no attention to the passive agent"
RPolite	"Participant was polite"
RReducePresence	"Felt reduced presence"
RStrongerResponse2Active	"Felt a stronger response to the active agent"
RSurprise	"Felt surprised"
RTaskOriented	"Participant was task oriented"
RTreatedUnreal	"Participant treated agents as unreal"
RTried2Engage	"Participants tried to engage with the agents"
RTried2EngagePassive	"Participants tried to engage with passive agent"
SReportedRealistic	"Self reported realistic behaviour and responses"
SReportedUnrealistic	"Self reported unrealistic behaviour and responses"
UApproach	"Participant approached the agents"
USpeak	"Participant spoke to the agents"

F.4 Network of Codes

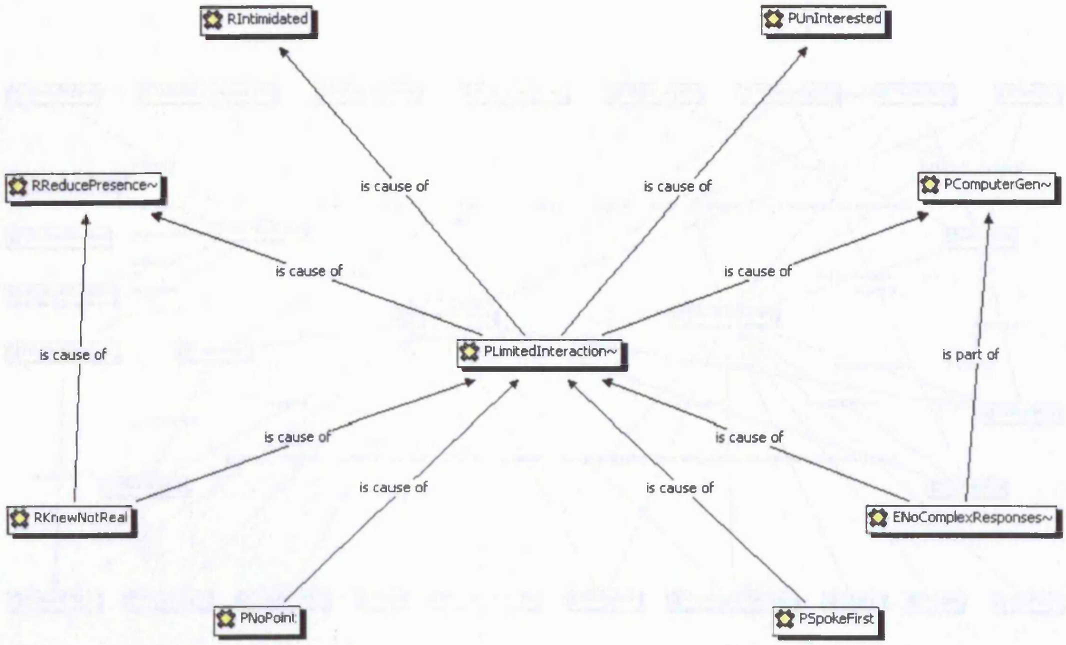


Figure F.3: Limited participant-agent interaction

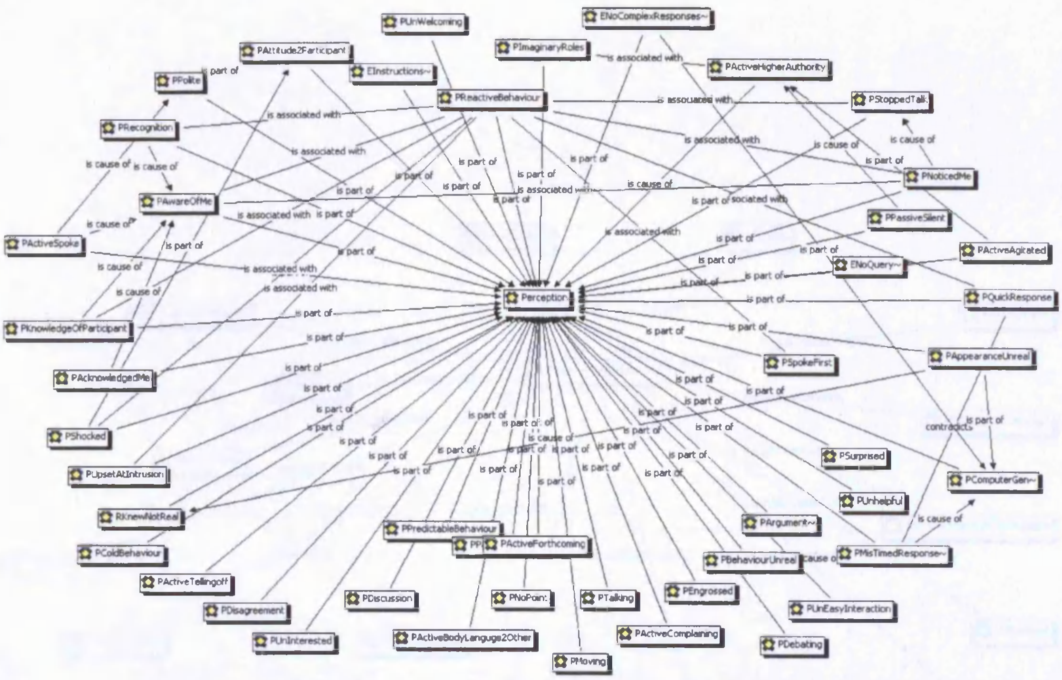


Figure F.4: Participant perceptions of the agents

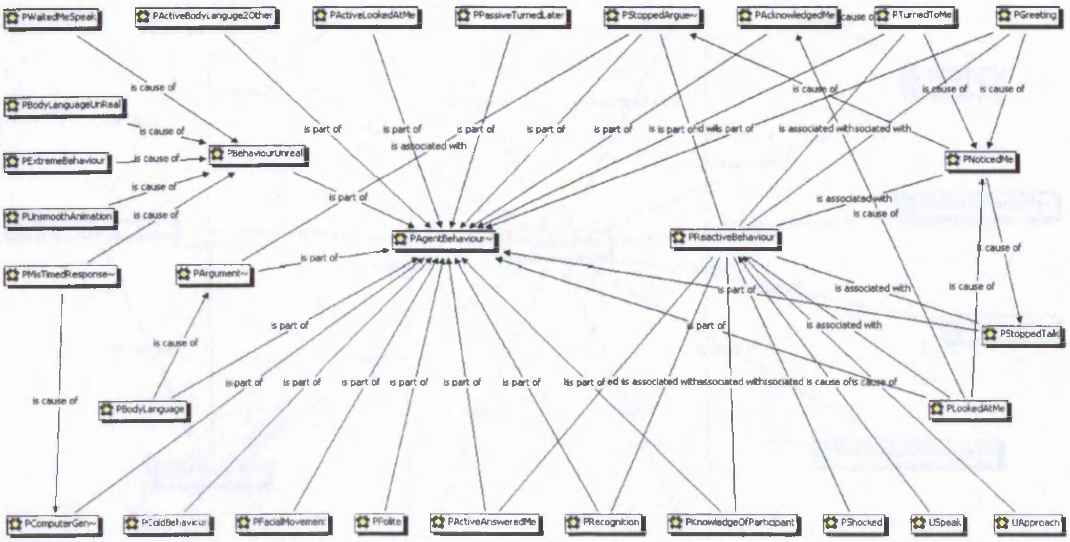


Figure F.5: Active agent's perceived behaviour

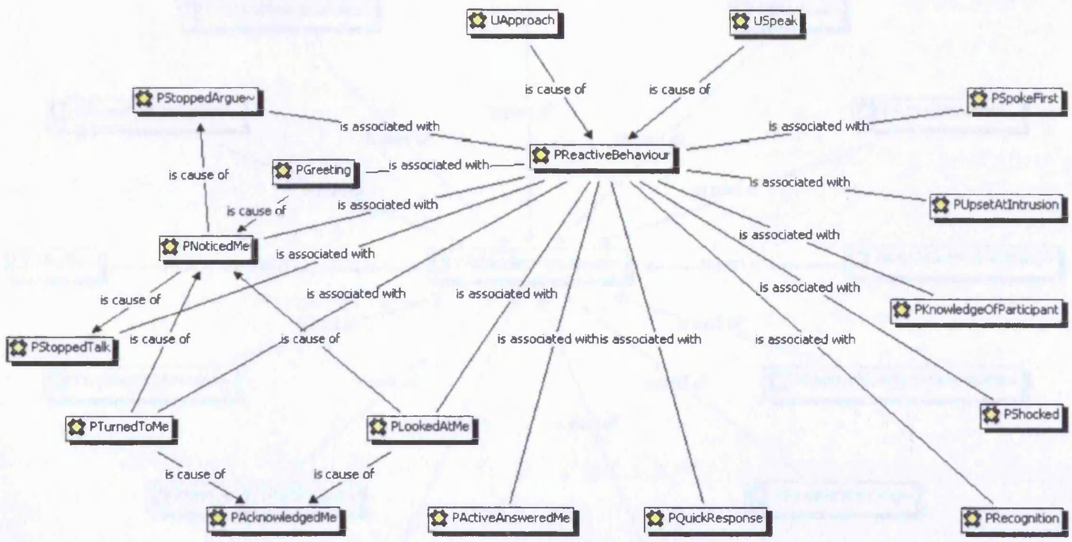


Figure F.6: Active agent's reactive behaviour

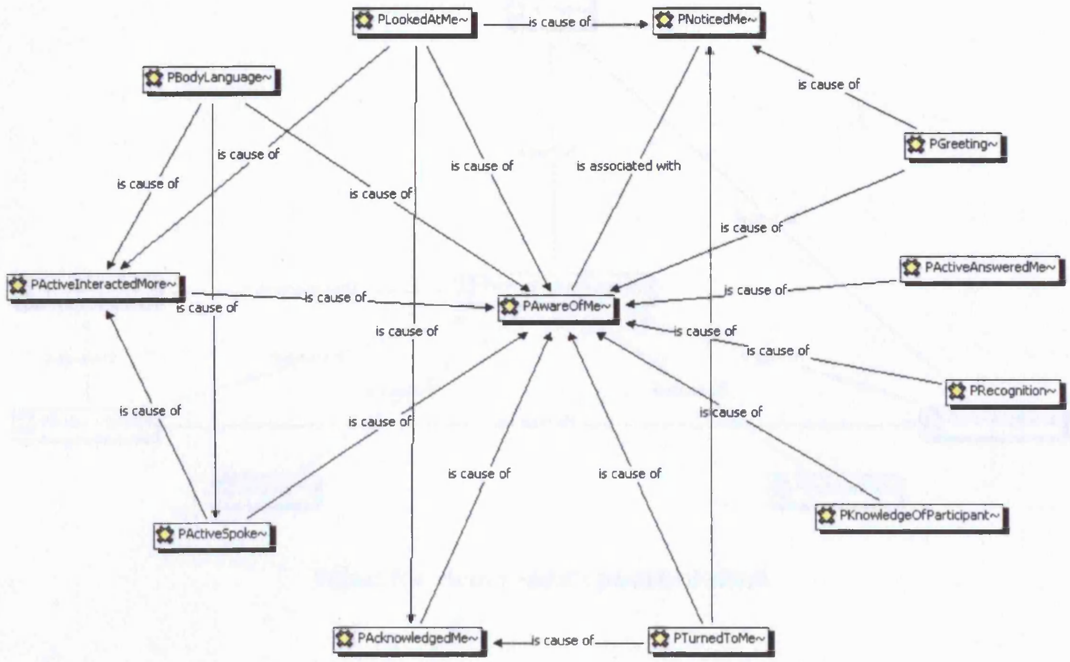


Figure F.7: Active agent's perceived awareness

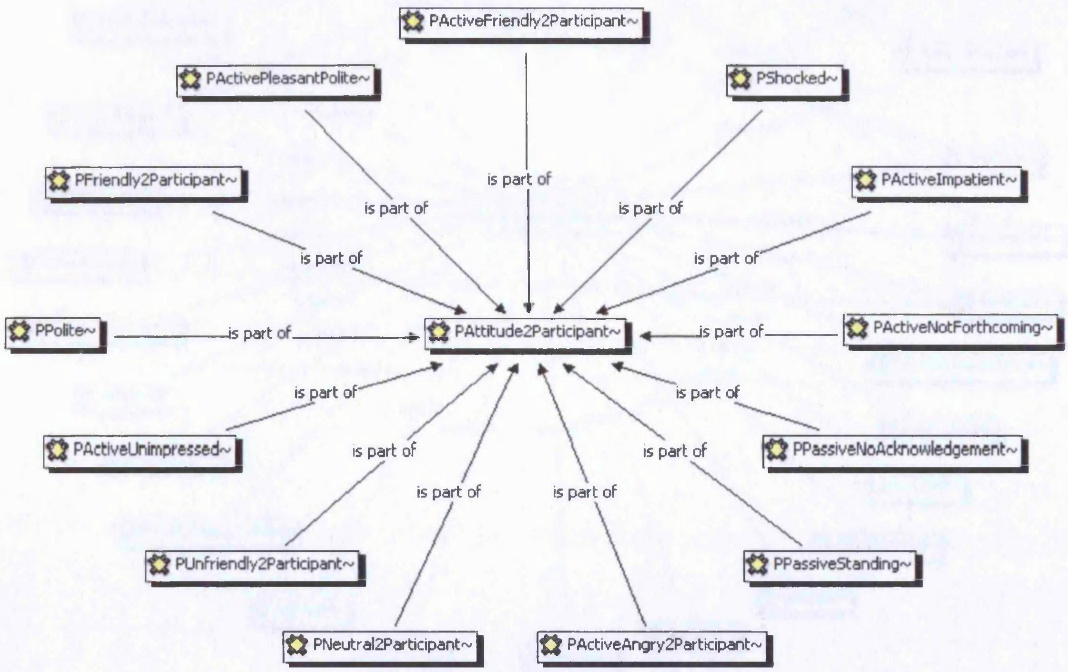


Figure F.8: Active agent's perceived attitude to the participants

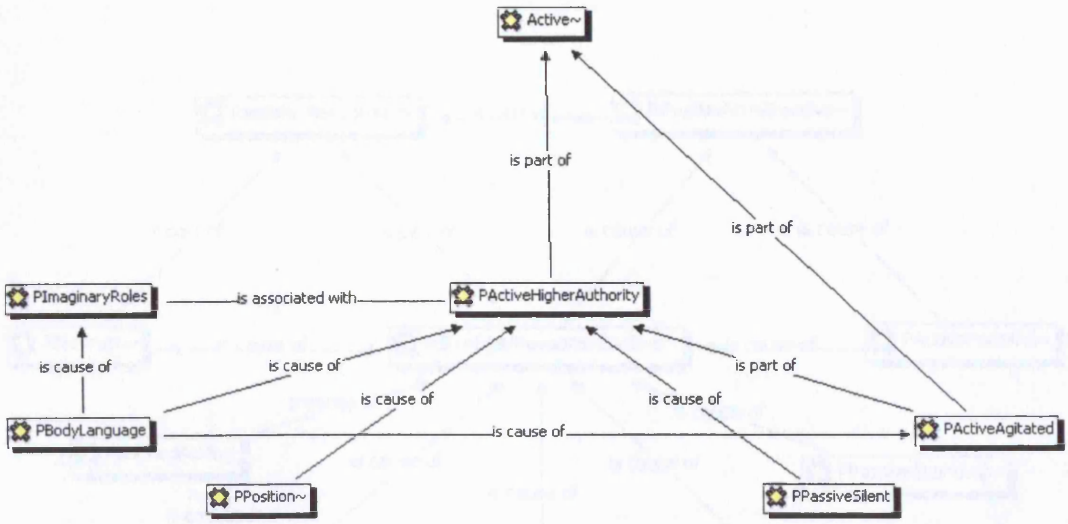


Figure F.9: Active agent's perceived status

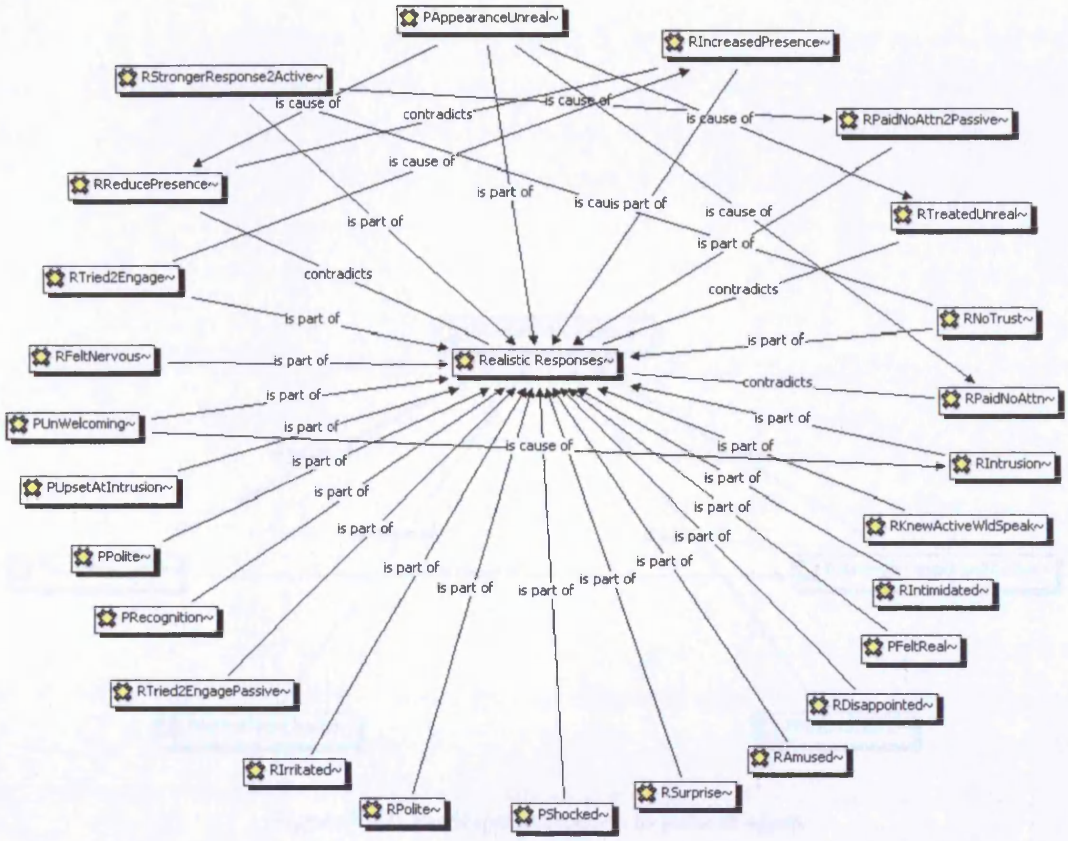


Figure F.10: Participant responses

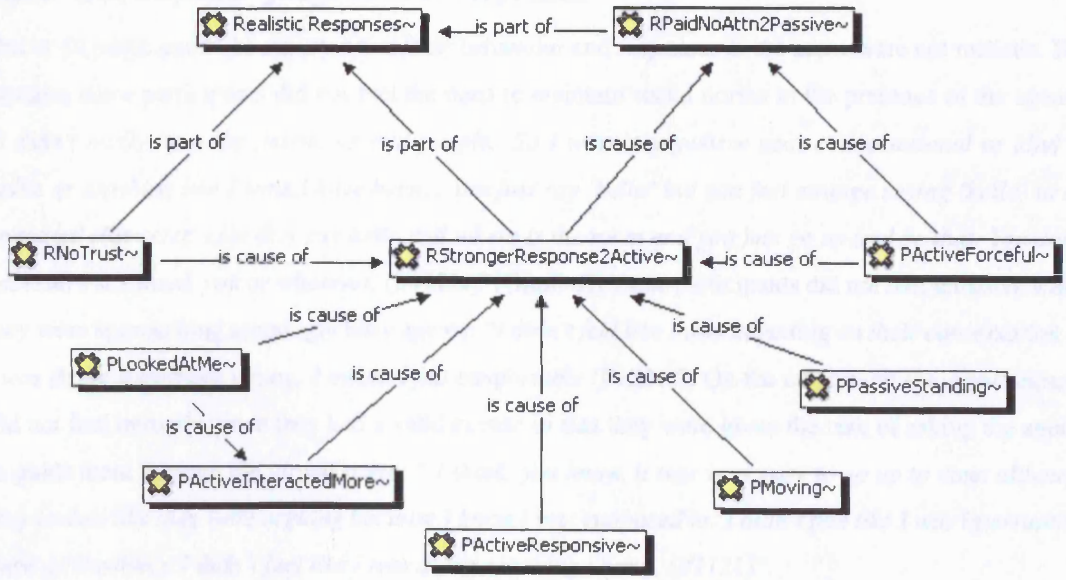


Figure F.11: Participant response to active agent

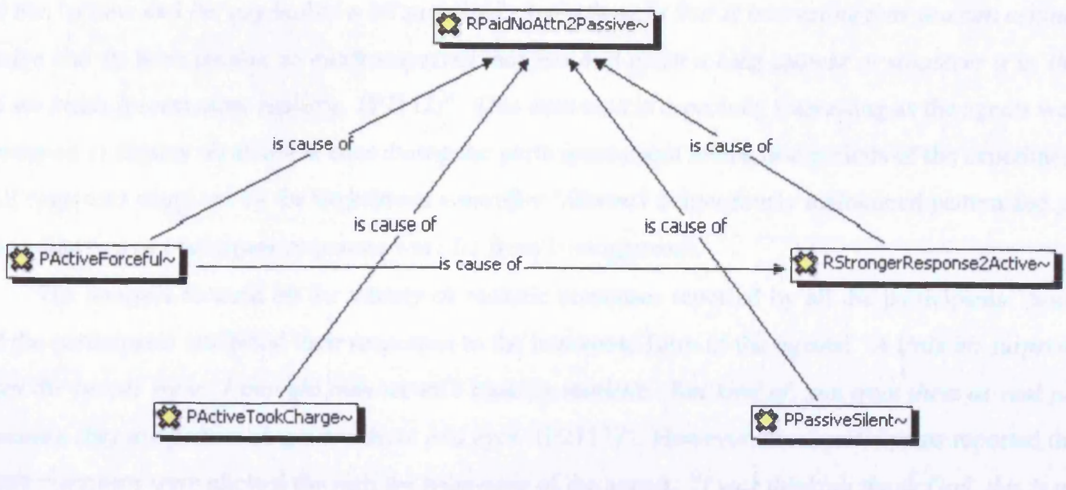


Figure F.12: Participant response to passive agent

F.5 Participant's findings from interviews

This section presents a preliminary version of the interview analysis with more comments.

F.5.1 Participants' judgement of responses

Out of 49 participants, 11 reported that their behaviour and responses in the maze were not realistic. For instance some participants did not feel the need to maintain social norms in the presence of the agents: *"I didn't really treat the people as real people. So I wasn't anywhere near conversational or kind of polite or anything like I would have been... You just say 'hello' but you feel strange saying 'hello' to an animated character. Like they say hello and where is the room and you just go up and do that. You don't necessary say thank you or whatever.* (P1122r)". Similarly these participants did not feel intrusive while they were approaching seemingly busy agents: *"I didn't feel like I was intruding on their conversation or I was doing something wrong. I mean, I felt comfortable* (P1221)". On the other hand some participants did not feel intrusive since they had a valid excuse in that they were given the task of asking the agents to guide them through the virtual maze: " *I think, you know, it was very easy to go up to them although they looked like they were arguing because I knew I was supposed to. I didn't feel like I was interrupting them or anything. I didn't feel like I was doing anything wrong.* (P7121)".

About half the number of participants (23) reported that their responses and behaviour in the virtual maze was surprisingly realistic while 13 participants reported experiencing a mixture of both realistic and unrealistic responses during the experiment. During the analysis of the semi-structured interviews, it became clear that the interpretation of behavioural cues in immersive virtual environment depends heavily on the character of the participant (interpreter). Some participants interpreted the behaviour of the agents in unexpected ways. For instance one participant reported that the active agent was surprised at the detailed answer the participant had given during one of the participant-agent interactions in the experiment: *"I was asked to say what was in the room, the expression to what I had said was different to the 1st time and the guy looked a bit surprised. And I thought that is interesting that you can actually judge that its been similar, so much surprised that you had given a long answer or whatever it is, that to me made it seem more realistic.* (P7112)". This comment is especially interesting as the agents were designed to display no affective cues during the participant-agent interaction periods of the experiment. All responses triggered by the experiment controller followed a rigourously maintained pattern and yet the collection of participant responses were far from homogeneous.

The analysis focused on the variety of realistic responses reported by all the participants. Some of the participants attributed their responses to the humanoid form of the agents: *"A little bit surprised that the people were. I thought they weren't visually realistic. You kind of, you treat them as real just because they are person shape and have two eyes.* (P2112)". However, most participants reported that their responses were elicited through the behaviour of the agents: *"I was thinking thank God, this is not a real environment because you know, the way they were going at each other, I would probably rather not approach them at the moment. So that seemed quite; how do you say? Realistic.* (P3121)". Some participants compared their responses in the virtual maze to similar events in the physical world: *"I think I guess it was fairly realistic, yeah because I remember one of the characters even saying 'pardon me', or* *something like, maybe he didn't actually hear what I said or maybe he recognised? that was surprisingly real because I am foreigner, people don't sometimes get what I say.* (P4211)".

F.5.2 Impressions of the interaction between the agents

Regardless of experiment condition, participants reported that their judgements of the interaction between the agents were made based on two cues: the body language of the agents and the audio properties of the verbal interaction between the agents. Generally participants reported the interaction between the agents as either a conversation/discussion or an argument/disagreement. In either case, the participants reported that the agents looked engrossed in their conversation since the agents only responded to the participants after the participants were within a certain distance:

- *The way they were talking seemed they were looked quite involved in the conversation.* (P1212)
- *It felt like they hadn't seen me and only when I kind of, got within couple of metres, they kind of responded to me.* (P1221)
- *They seem to be busy with themselves because they didn't even like look at me or anything when I walked into that room. They obviously didn't notice that I was there until I was literally right in front of them.* (P3121)

Out of the 15 participants who reported that the agents were involved in an argument, 13 were in the conditions where the active agent displayed angry postural cues towards the passive agent. These participants used body language as the main cue for their interpretation:

- *His (active agent) body language was a bit aggressive. Not with me but with the other guy (passive agent). You can see him (active agent) bending and shaking his hands, his arms actually and he was bending, he must have been desperate. The other guy (passive agent) was just assuming all the blame, just nodding his head. Because you don't really know what they were talking or arguing about but the attitude was one (passive agent) of submission and the other one (active agent) of aggression.* (P2111r)
- *They were kind of arguing because there was one (passive agent) staying still and the other one (active agent) was very projected at him - body language.* (P3111)
- *They looked almost like one of them was fighting, the other one (passive agent) just stood there.* (P5111)
- *On the rector scale of arguments, I guess it was about a 7. He (active agent) was facing the other guy and his body language and hand movements and he was concerned that something wasn't quite right and he wasn't happy about something. So that is pretty much the indication that an argument was going on. The other guy (passive agent) didn't seem to be responding, so perhaps it was just a one sided argument?* (P6121)

Other terms participants used to describe the interaction between the agents were disagreement, negative or heated conversation and confidential discussion:

- *I got the impression that it was a serious conversation of some sort. I would say more likely to be negative.* (P1112)
- *One of them was discussing his problems and the other guy was being the sounding board. So thats the left (passive agent) guy being the sounding board and the right one (active agent) being the one who is complaining.* (P1121)
- *I thought possibly they were having a discussion about some business really. Something was going wrong and the guy (active agent) seemed a bit panicky about that.* (P1221)
- *Some incident had happened probably, there was some past tense in there but... evaluation probably... hard to guess whether positive. I would probably guess slightly critical evaluation.* (P4112)
- *I don't think he (active agent) was angry, he was just maybe letting out? Like complaining?* (P6211)

Some participants reported hearing part of the audio conversation between the agents and using only that to form impressions of the agents:

- *My main observation of the people was that their conversation seemed a bit negative. Because they always seemed to be saying something like 'no no no no no that cant possibly be right'.* (P1211)
- *I think they were fighting. They were being a bit loud, I could hear them when I was approaching.* (P3112)
- *I guess it was just the tone. I don't remember what words they were using but it sounded there was kind of. They were definitely talking about something and it wasn't just kind of normal.* (P6212)
- *I kind of realised that they were having a conversation and started tuning a bit more to what they were saying and it didn't sound like they were having a particularly happy conversation, they sounded to me like something had gone wrong and they were discussing it and they were almost arguing about it. Generally the way that he (passive agent) was very quiet would say to me that he wasn't comfortable with the situation he was in and the other one (active agent) was very forthright.* (P7221)

F.5.3 Awareness and Reactive Behaviour

A secondary issue dealt with in one of the post-experiment questionnaire (Appendix E.5), partially to check if the agent behaviours were portrayed appropriately, was the responsiveness of the agents. Participants were asked to indicate if they thought each agent were a) responsive to them and b) interacted with them. Figure 5.17 shows that all the participants felt that the active agent was responsive to them and interacted with them. Interestingly a few participants felt that the passive agent was responsive and even more surprisingly three participants felt that the passive agent was interactive.

Participants reported that the agents were aware of the participant's presence and responded appropriately. This perception was mainly due to the participant's interpretation of the agents' behaviour but also in certain circumstances when the active agent appeared to have knowledge of either the participant's experience or the spatial properties of the virtual maze (Figure F.7). This effect was especially pronounced in the case of the active agent. The participant's response to the questionnaire item was used during the interview process to uncover the aspects of the agents' behaviour that led to the participant's perception that the active agent was more responsive than the passive agent.

One of the most significant behavioural cue that participants reported as being characteristic of perceived awareness was gaze (eye contact). Perceived awareness was associated with participant reports of the agents's perceived ability to notice the participant's presence by turning towards them and responding to their queries while looking at them:

- *They were looking at me and not looking somewhere else.* (P1111)
- *When I was walking up to them. They kind of ignored me when I was around but about a metre away, they looked at me. They both looked at me, they both faced me as if to say, 'he is here'. Yeah they were aware of me till the moment they sent me off.* (P1121)
- *Spatially they seem to be aware of me arriving and they sort of looking the right way, I mean they knew where I was, knew I was there. When I walked over, he sort of stopped, the guy in the suit on the right (active agent), stopped what he was doing and turned round and noticed my arrival and addressed me.* (P4122)
- *I approached them they obviously saw me, so they knew I was there, they knew I was coming.* (P6121)
- *Well he asked me what I had seen and how I was and also he looked very straight.* (P6122)
- *Purely by when you walked to them and they turned around, both facing you and in a way, almost introduce themselves or ask 'how can we help'. That definitely was a sense that 'ah, ok I am now in the presence of these two people'.* (P7112)
- *Whenever I walked up to them, they turned around quite quickly, they both appeared to be looking at me and the guy on the right said 'hello' to me as soon as I walked up.* (P7221)

Some participants were more certain about the perceived awareness of the active agent only either because they paid more attention to the more interactive active agent: "*When I approached, don't know about the guy on the left hand side - the more passive one, but the more assertive one did face me, so he was definitely aware that I was there.* (P6121)" or because the passive agent did not engage them in an active interaction: "*I think the right person (active agent) in the suit was aware, because he engaged in the conversation. I don't, I am thinking the left person (passive agent) was aware as well since he turned towards me but he didn't make any effort to converse.* (P7111)".

However as expected, nearly all the participants reported that the active agent was more responsive and therefore focussed mainly on the active agent:

- *I didn't make it a point but he seemed to respond to everything I said, so I didn't ask the guy on the left (passive agent).* (P1112)
- *The person towards the right (active agent) was more responsive. And probably more helpful.* (P3212)
- *I just totally concentrated on the guy on my right (active agent) because he responded. I felt like he was looking at me and asking me questions, saying you know, 'oh you and so on' and stuff like that.* (P5112)
- *I might be wrong but my impression was that responsive one, so probably the next one, I sort of more focussed on him.* (P5121)
- *I felt that it was just one of the two people responding to me and I feel fairly certain that it was the one on the right (active agent) but why I say that, why I feel that, I can't put my finger on it.* (P6221)

Participants also reported that the agents had a sense of awareness when the agents appeared to have some knowledge pertaining to either the spatial properties of the virtual maze or the participant's experience in the virtual maze:

- *They acknowledged that I have been there before when I came back from the first room by asking what I had seen. So then by telling me that, oh you need to turn around for that one because it is behind you, so yeah they did seem aware.* (P1211)
- *He said, 'try door...', cant remember what number it was. That was fine but I had to find the door and most people in a normal situation would go 'oh, its over there'. When he (active agent) said it was behind me, it made it more realistic because he was referring to the layout of the room.* (P1221)
- *I would say especially after the 3rd time, when I was looking for the door and I was told to turn behind me, so I assume that they must have been aware of my presence because otherwise it would be difficult for them to know.* (P5211)
- *He responded to what I said exactly. So it wasn't that generic, I was surprised actually that it happened. When I asked him (active agent) where to go, he responded by giving me the direction. He said just right behind you* (P5212)

F.5.4 Limited interaction: Due to agents or social context?

In addition to the visual appearance of the agents, the limited vocal responses and interaction capabilities offered by the agents was cited as the most significant reason for not responding to the agents as if they were real. The level of interaction between the participant and the agent depended on the participant's expectations. Each participant had various expectations of how the agents would interact with them. Some participants did not expect any queries or complex responses from the agents. Therefore these participants did not attempt to engage the agents in conversation since they felt it would be fruitless: "*In the back of my mind, I was aware that they were only computer generated. They have limited responses, so there was no point in trying to have a real conversation.* (P1111)". Some participants did not attempt to engage the agents in an interaction initially because they had no expectations of the agent's abilities. These participants grew more comfortable with the interactions after they had gauged the agent's abilities through experience: "*I approached them like I would do to real people, in that the role was to go and talk to them and get directions. so I would treat them as real anyway. I guess the thing is, by doing it the first time, you get a response from them because I wasn't sure that they would actually respond to anything I said. So I wasn't sure but when they did respond to my normal question then the 2nd time I was more comfortable in talking to them in normal phases and so on and getting their response and carrying on, so that was good.* (P7112)". When the agents respond in a manner that surpassed the participant's expectations, participants attributed a sense of realism to the agents: "*They caught me off guard a couple of times because they asked me what I had seen in the room, and I didn't anticipate them in doing that. So that brought an extra bit of realism to them.* (P1211)".

Participants were surprised to experience responses to the agents especially when the participant had a preset expectation of the agents' ability: "*One thing is exactly like I expected which is like you know graphically they just won't look like real people, and therefore, you will find it hard to take them seriously as real people. But the fact that they were talking and seem a bit negative... that's a bit weird... was a real sort of bad vibes going on there. Which you know bad vibes from a bunch of graphics. That was maybe a bit more than I expected.* (P3121)". However, a few participants expected the agents to be highly informed despite their visual appearance: "*I had the impression that when I walked in there, they would be a bit more helpful in terms of telling me what to do as opposed to me asking and they seemed to only react if I said something.* (P1112)". This impression might have been developed after the participant was told that the agents would act as guides to the virtual maze. Some participants expected to conduct a more naturalistic and interactive conversation with the agents: "*It was a more like a automatic relation, you know. I ask for information, he provided the information. It was a more descriptive relation, ok, you have to do that, and I was waiting for the information to continue.* (P2121)".

After unsuccessfully trying to engage the agents in a more complex interaction, participants often resigned to simply getting the information required to complete their task:

- *I think this time not so much to interact but just to get the information I needed because I realised they weren't particularly interested in talking or interacting.* (P4121r)

- *My chances with the people were very limited and I didn't ask their names or anything else because it didn't seem to matter* (P5122)

- *I think I initiated something like what do you do for living, obviously he wasn't. I don't know but obviously I was detecting he wasn't going to continue my approach initiating any interactivity... I think I asked him and he just said room 4 is good or go to room 4. I thought yeah this guy is not got a lot of personality.* (P6121)

Participants were especially attuned to timing errors in the active agent's responses either due to network lags or human errors:

- *There is a sort of delay in conversation. They are not really interacting with you. They have just got this set of responses.* (P1111)
- *The speech was a bit halted, I think. Maybe the computer switched to something else or I don't know what happened but it is like something was selecting what he needed to say instead of it actually coming from him.* (P1221)
- *In a real conversation, you find people talk pretty much to one another, like one person say something and the other person starts before the next person finishes, which you didn't have that.* (P3121)
- *The gaps between the things I said and the things they said weren't quite right. Sometimes they were slightly too short and sometimes slightly too long. And so like the last time, I started talking, he interrupted with an irreverent answer almost straight away.* (P4112)
- *Their response wasn't as quick as if it would be a real situation. In a real situation, you interrupt someone and you ask. And the other person, as soon as you finish your question, they respond. Not necessary with an answer, they may say 'I don't know' but they respond. Here there was an interval. There was a delay from the moment you stop talking to the moment they start doing the talking... It is not spontaneous, looks a bit artificial, looks like a bit computerised.* (P5121)
- *The guy on the right (active agent) said hello first and started talking to me. In fact he had a tendency to you know, kind of start talking when I was talking as well. Which would be rude, if it was a real person but it was just, you know the way their responses were.* (P7221)

On the other hand, some participants reasoned that the limited interaction was either due to the social context within which the interaction took place, a lack of the agent's interest in the participant or the agent's previous emotional state:

- *He didn't seem too happy. They didn't seem friendly in terms of. ask me why I am here or anything, they didn't seem to want to have a conversation, just get on with it.* (P1112)
- *It was quite an uneasy meeting of people because they were uneasy before I got there. I got that impression and they were fairly brief. So I guess they weren't really interested in me.* (P2112)
- *The one guy's behaviour towards me was pretty neutral typical, I don't think he (passive agent) liked me very much.* (P3112)
- *They asked me what I saw and when I told them, they didn't seem very interested in the details. So I don't know why they asked, they weren't interested.* (P4121r)
- *I had the task to do and the way they responded told me that they knew that I had the task to do. And they didn't try any small talk or anything like that, they just told me, go look in the room immediately, so I just went.* (P4221)

- *I walked into the central room to get advice to walk to which room, since my purpose is only to walk to which room and not looking for friends, so I only seek advice from the person in black (active agent).* (P6112)
- *I got more of a response out of the right guy (active agent), he seemed like he was you know quite confident. They both seemed pretty disinterested. Pretty kind of, you know, focussed on themselves.* (P6212)

A significant number of participants reported that the limited interaction was a result of the task given to them. Since the task involved using the agents as guides to explore the virtual maze, the participants paid very little conscious attention to interacting with the agents:

- *I didn't really interact with them, didn't try to engage in the conversation, didn't at all try and talk to them. The guy on the left (passive agent). So it could be part of me not being very social or it could be part of that I was thinking that I should just get information, that is it... I didn't think it was necessary since I already have the information that I wanted.* (P3112)
- *I think I was more driven by the mission, that is know how many dust bins in each room, so I wasn't actually interested in making a conversation with them.* (P3211)
- *I interacted with them quite briefly all three times. If I walked up to a couple of security guys at a airport and they were having a conversation and I wanted to know which door to get to my gate, I would try and find out. If one of them responded to me and looked like he was the one who was expecting to answer the question. Then... I mean, if it was a party or something like that, you know somebody I just met at work, then I would try and find out who the other person was but not when I had a specific question.* (P4112)
- *Well there was an assigned task and they told me what I had to do and that was all the kind of interaction that I was interested in.* (P5221)
- *I didn't really care, I didn't feel like they could really see me, it didn't matter how I acted or I spoke to them, it was just a matter of getting instructions. Although it did seem quite real, it is not that I didn't care because I didn't feel real that I was there. I guess I just didn't feel like it mattered how I acted socially.* (P6212)
- *I did not feel the need to, it may be by the task in the sense that I felt, I am just here to get the directions from these 2 people, there is not much need in interacting any further.* (P7112)

F.5.5 Participant responses to the agents

The agents were designed to have a neutral attitude towards the participants, however, the qualitative analysis of the interviews uncovered that participants had varying impressions of the agents' attitude towards them. Over 45% of the participants reported that the agents had a formal or neutral attitude towards them with slight variations. In keeping with their design, the agents' behaviour was described as polite: *"They were I think cordial and polite sort of thing. They showed a bit of shock when I arrived* *but they were polite and they answered my questions and they seemed quite neutral towards me. They didn't sort of enquire to why I was there or anything like that. They simply told me what I needed to know* (P1211)"; slightly upset at being interrupted but still neutral: "*The guy when I walked up to him kind of broke up the conversation. Pretty much immediately when I walked up and turned around... Possibly like slightly upset at the intrusion I guess. Maybe. But the way he spoke to me was pretty much neutral... He didn't seem annoyed to talk to me but he didn't really seem like happy to see me or anything like that.* (P1212)"; friendly but cold (perhaps formal): "*It was a friendly behaviour... Very friendly but cold behaviour. Yes, an English behaviour.* (P2121)"; sufficiently accommodating: "*When I spoke to him, he wasn't particularly angry or rude to me, he wasn't overly pleasant but he seemed to able to sort of leave his things so that he could talk to me for a little while* (P4122)"; instructional: "*Neutral, I mean they weren't very emotional in any way, they just came up, they saw me, they responded, they told me what to do.* (P4221)"; in keeping with the situation: "*I didn't sense over positive, over friendly or negative... It was ok for the situation.* (P5121)"; dutiful: *I don't think they were particularly happy to see me, I don't think they were particularly annoyed to see me, I think their job is to be polite and point me to a room.* (P6121) and generally behaved as can be expected: "*Fairly neutral, just what I would expect if I was just a stranger going up to 2 people and asking them for directions.* (7221)".

Other participants reported different attitudes for each agent since the active agent was more interactive than the passive agent. Even though the only difference between the behavioural cues displayed by the agents towards the participants was in the level of subtlety, the overall effect seems to have created a noticeable difference between the apparent personalities of the agents. Participants interpreted the behaviours of the agents in a holistic manner by comparing the active agent with the passive agent in parallel:

- *It just that he (active agent) seemed more engaged with me and the other person just seemed to be standing there* (P1111)

- *The guy on the right (active agent), I thought he was fairly neutral or maybe a little bit more friendly. The guy on the left (passive agent), I thought, I don't know, if he was scared to talk.* (P2112)

- *I think the person on my right hand side (active agent) is nice, the other person (passive agent) has no apparent... I didn't see him because he never respond me and I don't know* (P3122)

- *I thought he (active agent) was kind of a warm person in a computer way and the other person (passive agent) was very cold. He just didn't actually interact at all.* (P5212)

- *The gentleman on the right is kind and gave instructions to me and then the other one did not have any response to me.* (P6112)

Furthermore some participants perceived a change in the agent's attitude even though there was none:

- *I think the first two meetings, they were quite similar you know, it was kind of a negative atmosphere. But then the third meeting, it seemed to have been resolved and he (active agent) was a bit more kind of happier to tell me which room to go to.* (P1221)

- *He seemed helpful at the beginning but he got a bit impatient about me going back to ask 'where to go next'. The guy seemed pissed off in the end.* (P1112)

- *I think they were positive in the beginning and after that the right guy (active agent) started to complain about the left guy (passive agent) because of something he had done.* (P4111)

- *The first time the person I talked to (active agent) looked kind of disturbed... I didn't think they really wanted to start the conversation but the 2nd time, it was more friendly whereas the 1st time was more direct.* (P5212)

Feelings of intrusion: One of the most significant responses reported by participants was the feeling of intrusion when interrupting the agents at the end of the approach periods of the experiment. This was especially the case when the agents were perceived to be involved in an argument or disagreement with each other. Participants reported feeling hesitant, embarrassed and uncomfortable upon intruding on the agents' conversation:

- *I felt they were really arguing with each other. The right one (active agent) was like shouting at the other one. And then I got there, yeah it was interesting, I thought it was pretty real and I felt like I was interrupting something... I felt like embarrassed. Should I be here?* (P4111)

- *I felt that since they had seemed like sort of argument, at least it had some emotion with it... these guys are having some kind of heated conversation and its in that situation, its difficult to go and talk to them, go and ask them. You don't want to bother people when they are engaged in heated or intense conversation.* (P4211)

- *I thought I was intruding on their private conversation and I wasn't really aware of the person on the left (passive agent). He basically ignored me but the person on the right, yes I thought he was a real person* (P5122)

- *I was holding back a bit if I could ever hear their conversation or to see if it would finish, if it would be any conclusion but the assertive guy, the boss (active agent) was still not happy and the less assertive listener (passive agent) didn't seem to be reacting back.* (P6121)

- *Well one was very friendly and forthcoming (active agent), a bit formal, a bit you know like an administrator but still friendly... and the other one (passive agent)... I think I may have intrusion. So the conversation they were having was important for him (passive agent) but not so much for this (active agent).* (P6122)

- *I walked up to the 2 people that seemed to have been in an argument. And since I didn't want to interfere with the argument, I just asked them that I had a task and I needed help. The main thing* *was that I saw them in a conflict, some type of argument, so I held back interacting with them. Similarly if I saw two people fighting in the real world, I wont approach the person anymore... I would choose the one that weren't in conflict with each other... I kind of felt bad for the person on the left (passive agent).* (P7111)

- *I think if there had been other people around having a different conversation that wasn't as antagonistic as those people, I would have gone and talked to them instead.* (P7221)

In addition to feeling intrusive, one participant also reported feeling anxious: "*I didn't really want to get to know him because unknown persons makes me more anxious so... and I am already not very good understanding conversation* (P3112)".

Politeness: Another response that surprised some participants was their tendency to be polite towards the agents despite knowing that the agents weren't real: "*I didn't think they would be able to have a discussion back with me. So I anticipated a very limited capacity because they were generated by the computer, so I just sort of spoke to them 'hello, which room am I suppose to go into?'. I was aware that I was polite though because I kept saying 'thank you' after they told me where to go which obviously is unnecessary because they are not real.* (P1211)". Participants reported that this response was automatic: "*I found myself thanking him, for telling me where to go next. Which is what I would do for a real person. Which did feel slightly strange saying that to an image but I think that is how strongly how I respond to real people. I do tend to say please and thank you. So that is my natural dialog response. So yeah I was quite surprised as I found myself saying it. But it sometimes seems the right thing to say.* (P5122)". Some participants reasoned that this was because the agents were more realistic than they had expected: "*They were more realistic than I thought it might be, but you know still... not very convinced that they were real but I think I was being very polite to them and saying thank you and things like that. You know hello and thank you and all that.* (P7221)". One participant reported feeling hesitant in carrying out his plan to try and listen in on the agents' conversation since it would have been socially impolite: "*I was kind of curious to look at them but then I probably moved because it felt socially awkward to just... because you did have the sense that they were responding to you... It felt like if I just kept standing there and staring at them, it would have been kind of inappropriate.* (P6212)". This is particularly interesting since the participant's curiosity to test the agents was overcome by his desire to maintain social norms.

F.5.6 Participants' impressions of the agents' persona

Perceived authority and status: Most of the participants (39 out of 47) compared the interaction they observed between the agents with similar interactions they had witnessed in real life and assigned roles with higher authority to the active agent due to a combination of the perceived persona of the agents: "*He (active agent) was obviously the one dominating between the two of them and he looked more, I suppose authoritative because he had the jacket and smart* (P5111)"; and the attire of both agents: "*The kind of impression I got was of the guy in the suit (active agent) being the boss of the guy in the white t-shirt (passive agent) basically. He worked for him or something. Maybe that was to do with the suit.* (P7221)".

Participants also assigned a position of higher authority to the active agent due to the manner in which they interpreted the behaviour of the agents:

- *So there was definitely an argument, the way they were acting, the movements, the body language, over the guy with the black suit (active agent) and the other guy (passive agent) was very passive. The clothes was very different as well, so there was definitely a relation of power and there was a man in the suit and the other guy was wearing a street cloth or something. I think the guy with the suit and the attitude and the way he addressed the other guy made me think, he was, perhaps not the thing but the way he was speaking, I thought it was a working issue. This guy had some power over the other guy but not on, lets us paternal or emotional possibility. It was not his son or his relative or because of the way he was addressing him. So they were definitely speaking about work or business.* (P2111r)
- *I thought it was very realistic how they both turned to look at you but then only one person speaking to you... If the boss was talking to someone and someone else comes in, he would want to take control in the conversation, so that was good.* (P2112)
- *I think its probably the fact that there was no negotiation going on, I think friends tend to negotiate things even if they get angry, they tend to settle a negotiation. Whereas here, the role was one of the dominant and submission.* (P4121r)

Some participants went further than simply attributing the active agent with a higher status. The participants reported using their perception of the active agent's to choose the active agent for interaction since the active agent looked commanding: "*I think because of his position, he was wearing a suit (active agent) and he was like... just all the time looked like he was commanding the other one. Probably that is why I went to him* (P4111)"; was more engaging: "*I noticed two people in the distant talking, as I got nearer and they seized their conversation and one of them (active agent) took the initiative of facing me. Obviously he wanted the interaction because he was looking at me face to face, I didn't ask him anything.* (P6121)" and looked more approachable: "*I meet two people, people in blue shorts (passive agent) and another in suit (active agent). And the people in the suit looked more gentle, more older that people in blue short... I asked him which door should I come next.* (P7211)".

Furthermore two participants reported playing a passive role in their interactions with the active agent because they felt intimidated: "*He answered my question immediately... and waited for me either to go... I didn't dare to say anything more... I guess I was intimidated by this guy or also surprised perhaps. I just wanted to follow the task I was told. I complied to the guy in the suit attitude (active agent). I was very passive in that sense.* (P2111r)" and judged: "*I think the guy on the right (active agent) looked at me but I think he remained quite distant and I had a sense he was kind of looking down to me.* (P4121r)". One participant felt that the active agent was unimpressed with their performance in the maze: "*I got the impression of the guy in the suit (active agent) shaking his head at one point and waving his arm and very very human to go with the sort of anger sort of thing, you know and not being satisfy with the guy with the shorts (passive agent). They are very sort of human responses, I think seeing* *this kind of body language which is very human, that probably made him seem very real itself. He might have nodded his head a couple of times when I was talking to him and I guess a very human response was, he seemed very disappointed.* (P4122)".

Impressions of the passive agent: Participants attributed the passive agent's lack of responsiveness to non-technical constraints. This was especially the case when participants tried to interact with the passive agent and were unsuccessful:

- *He (passive agent) seemed less friendly than the guy on the right (active agent). He didn't make any effort to say anything to me at all.* (P1212)
- *They both turned and faced me, so they both acknowledged me but only one would speak to me and I looked at the guy on the left (passive agent) as well but he was not saying anything.* (P2112)
- *Well I mean I wasn't consciously thinking right I am going to be normal, I just went up and 'how are you doing guys, you alright there', you know. I said to one of them (passive agent) that I liked his shorts but he didn't respond. In fact he didn't respond at all actually. So I just left him. He seemed like he was in trouble anyway... I definitely got the impression that the guy on the left with the shorts was in trouble. I thought I heard directions being given before I got there but he definitely seemed in trouble and possibly scared of the other guy (active agent).* (P4122)
- *I think interestingly I sort of had the sense that he (passive agent) was kind of relieved that something had come along to interrupt his interrogation.* (P4121r)
- *I think the other one (passive agent) took a step back or at least sort of seemed to be less impulsive or less confrontational, I don't know.* (P7121)

Generally participants did not pay attention to the passive agent since the passive agent did not respond actively to the participant:

- *The other guy (passive agent), I think also turned to face me. But I wasn't really aware of what he was doing. Not sure if he was actually looking at me or I didn't think he ever said anything to me.* (P1212)
- *When I spoke with them, then I noticed the guy on the right (active agent) was a bit more responsive, spoke to me more than the other one (passive agent). Maybe the other one said something too but I cant remember any of it.* (P4211)
- *The guy on the left (passive agent), when I went up to them, I think he turned and they were both facing me but he wouldn't say anything. So I didn't even bother to look at his face.* (P5112)
- *I don't think I was really paying much attention to the guy on the left (active agent) because well he wasn't paying much attention to me.* (P7221)

Other participants did not pay attention to the passive agent because the active agent took control of the interaction: "*He (active agent) took over the situation immediately. But he doesn't allow you to* *see the other guy (passive agent) or what is going on. So it was a bit aggressive in that way because they immediately react... he immediately reacted to my presence. And he established his distance, so that I wouldn't interfere with what they were doing and I could go on with my business basically.* (P2111r)" and appeared to be the agent in charge before the participant started interacting with the agents: "*They turned towards me and definitely the person that I had the impression was talking, was facing straight towards me... I didn't really look much at the other person (passive agent) at all, I am not quite sure why... well he (active agent) seemed to be in charge I guessed because I got the impression that he was telling the other one person off.* (P7121)".

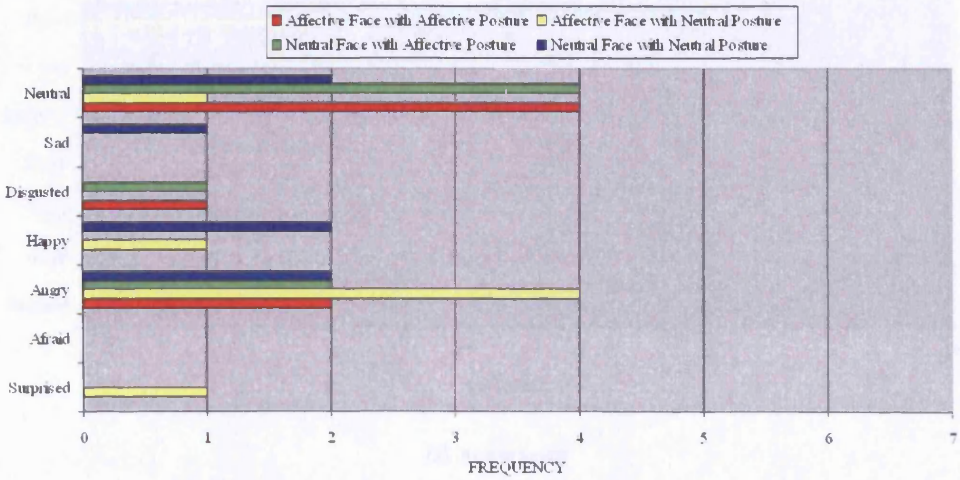
A number of participants did not interact with the passive agent because the lack of responsiveness made the agent appear indifferent: "*The left one (passive agent) never talked to me. Because he was not exactly neutral like I placed in the questionnaire but there was not the correct adjective, he was more like indifferent* (P4111)"; and cold: " *I didn't talk to him, interact with the other guy (passive agent) because he seemed colder than the other one... I think that both of them should have come to me and introduce me to the place... because he was the one I talked to and because the other one didn't engage, I didn't then decide to engage with him.* (P5212)". This effect was observed throughout to such an extent that participants could not accurately described the visual appearance of the passive agent. One participant did not even remember the gender of the passive agent:"*The person on the right hand (active agent) tells me what I had to know and they go back to their little thing and the person on the left (passive agent) might have said something, you know as in 'hi' or something like when I approached them but I don't think he or she gave me any directions at all.* (P3121)".

However the passive agent's behaviour was not judge as unrealistic. It was generally accepted that the behaviours portrayed by the passive agent for the role it played was natural: "*One was more helpful than the other but again, in a real situation that is the what to be expected, one is already helping and the other wouldn't and that seemed pretty natural... but yeah the one with the black jacket and trousers (active agent) was more helpful than the other one (passive agent).* (P7112)" and in keeping with the scenario: "*It was always the one on the right (active agent) which was addressing me and the one on the left (passive agent) never said anything. So I guess, in that way I was responding as if they were real people because I expect if two people were having an argument and one was dominant, and then you came up to him and asked, you know for directions or something, it is going to be the one who is dominant in the argument that is going to say something whereas the other one might go quiet and not feeling comfortable with the situation.* (P7221)".

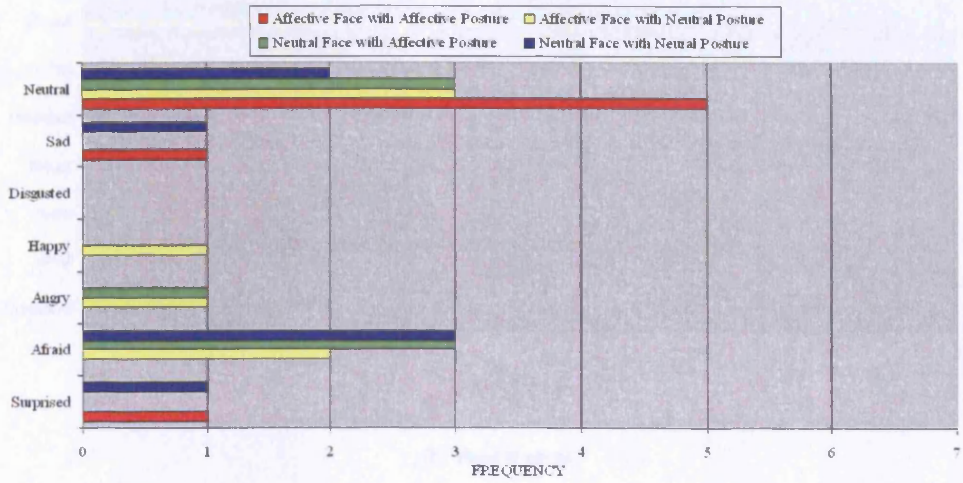
Agents were not inclusive: In keeping with participants' feelings of intrusion (Section 5.7.4.4), some participants perceived the agents as being unfriendly. These participants described the agents' behaviour as unhelpful and dismissive: "*When I realised that my presence wasn't particularly welcome and they weren't necessary going to give me an awful lot of information or be particularly helpful, I found myself quickly giving up on any social interaction. I didn't really want to engage these people in a conversation... I think sort of, the guy on the right (active agent) seemed vaguely tolerant but only because its something he had to do. But not very friendly and a bit dismissive. I think just the general* *diffidence of both characters, but especially the kind of, yeah the manner on which the character on the right managed to... I think interestingly a lot of what is expressed is subliminal, I can't really put into words exactly why I wouldn't want to interact with these people, but I guess its a lot to do with eye contact or its to do with being made to feel unwelcome or not receiving these any of these welcoming signals that I might otherwise get say among friends... None of the common courtesies were expressed. And when I offered information about what I had seen, I got a pretty neutral, it was just a strange, 'hmmm'. Rather than a positive 'oh really' or 'that is interesting'.* (P4121r)"; not welcoming: "*The people were not very welcoming. I think the real thing is that, I felt like they were in their own world, they knew what this was the space I didn't know and this is main reason why I didn't feel comfortable... the people were not very welcoming and so made me concentrate on what I had to do to finish and get out of there, even though I was not nervous. I could have spent more time if I was more in a comfortable situation.* (P2111r)"; and too engrossed in their interaction: "*They were already in a conversation that didn't seem very prone to outsiders... Every time I walked close to them, they stopped and turned around to me normally, so I got the feeling that they didn't want me to be involved in it at all.* (P7111)".

Since the active agent was more heavily involved in the interaction between the participant and agents, the active agent in particular was perceived as impatient: "*I think I would comment on the attitude of the guy on the right (active agent). He seemed pleasant at the beginning but I think he got a bit impatient as I asked for questions... especially when I didn't hear what he said.* (P1112)". One participant reported that the agents were trying to cut the interaction short since the agents were interested in continuing their previous conversation with each other: "*They were trying to get rid of me because they wanted to talk about... what they were talking about. I think that is why I said they were trying to get rid of me. So one guy (active agent) was instructing or telling the other guy (passive agent) off. And he was silent. It wasn't his place to talk to me. Because the other guy was taking control of the conversation and continued the conversation whenever I came, yeah he was probably... his face looked neutral but he was probably trying to get rid of me.* (P2112)". Another participant reported that the agents did not want to have an interaction with him initially but this changed during the experiment: "*The first time the person I talked to looked kind of disturbed, that I should, as I went to them. I didn't think they really wanted to start the conversation but the 2nd time, it was more like... more friendly* (P5212)".

F.6 Participant's judgement of agents' emotional states

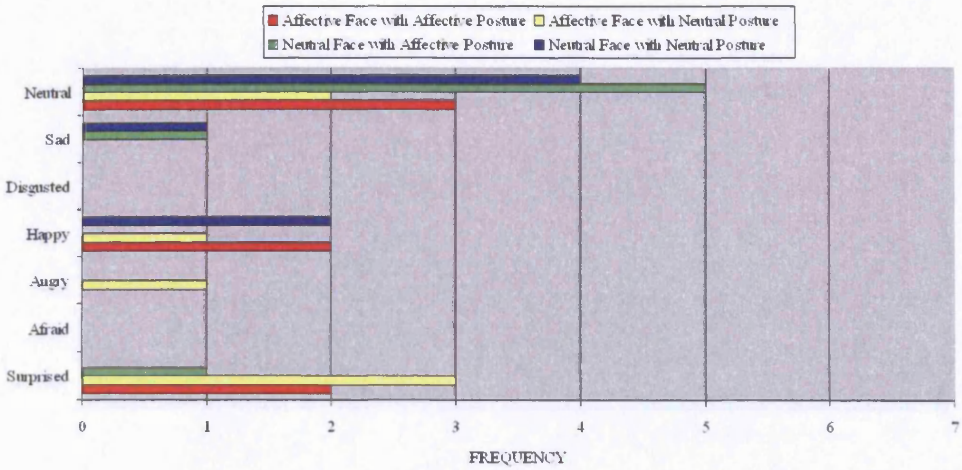


(a) Active agent

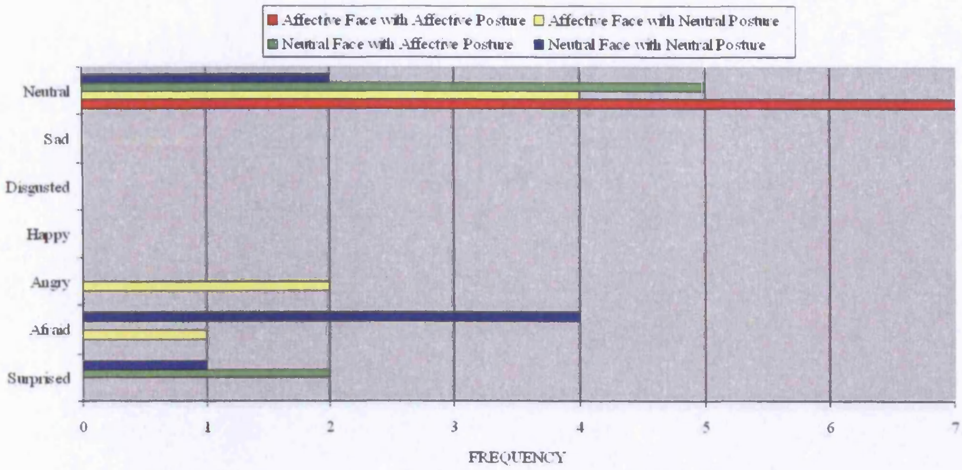


(b) Passive agent

Figure F.13: Participant's judgement of the agents' emotional states towards each other in the Sad conditions

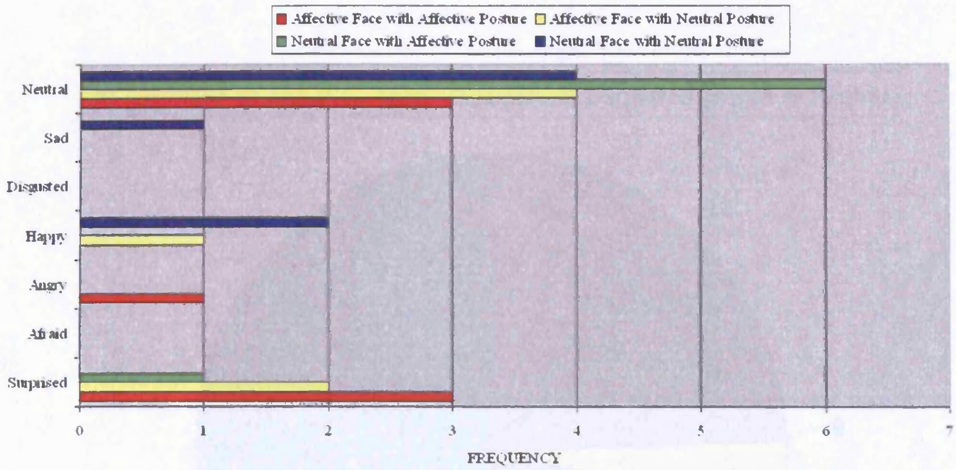


(a) Active agent

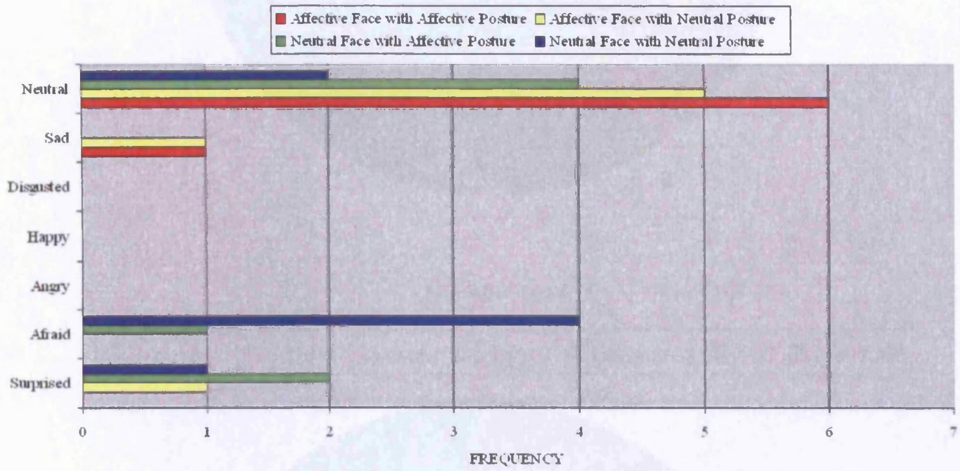


(b) Passive agent

Figure F.14: Participant's judgement of the agents' emotional states towards the participant in the Angry conditions

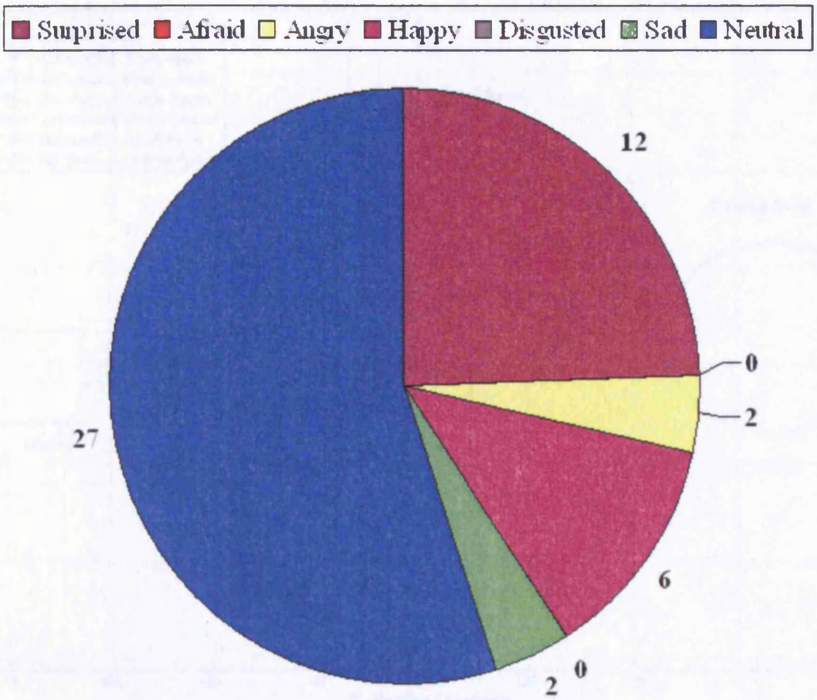


(a) Active agent

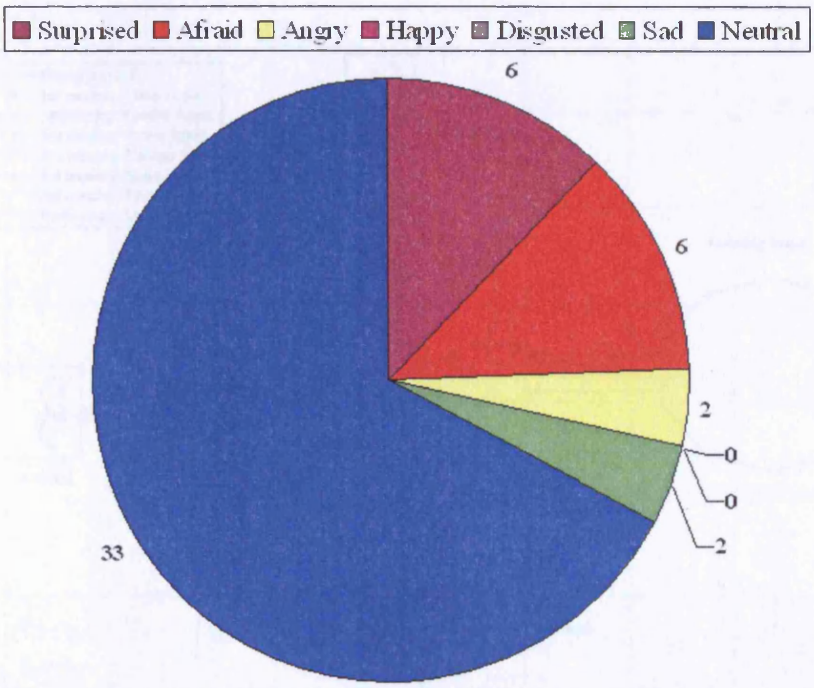


(b) Passive agent

Figure F.15: Participant's judgement of the agents' emotional states towards the participant in the Sad conditions



(a) Active agent



(b) Passive agent

Figure F.16: The participant's perception of the agents' emotional state toward the participant.

F.7 Tracking Data

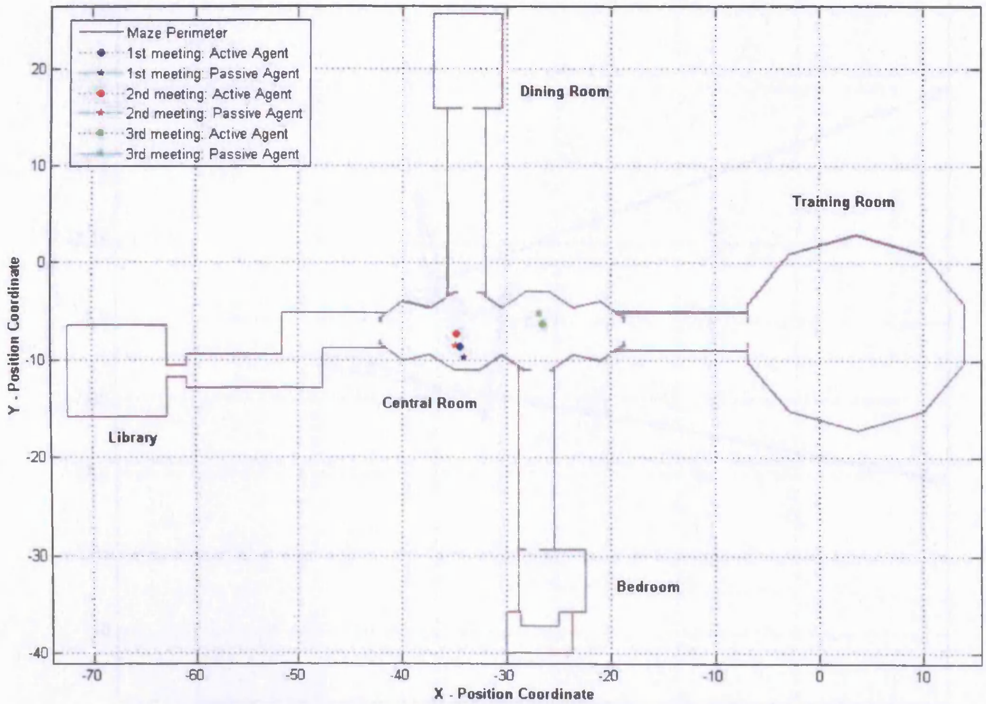


Figure F.17: Map of the maze using tracking data

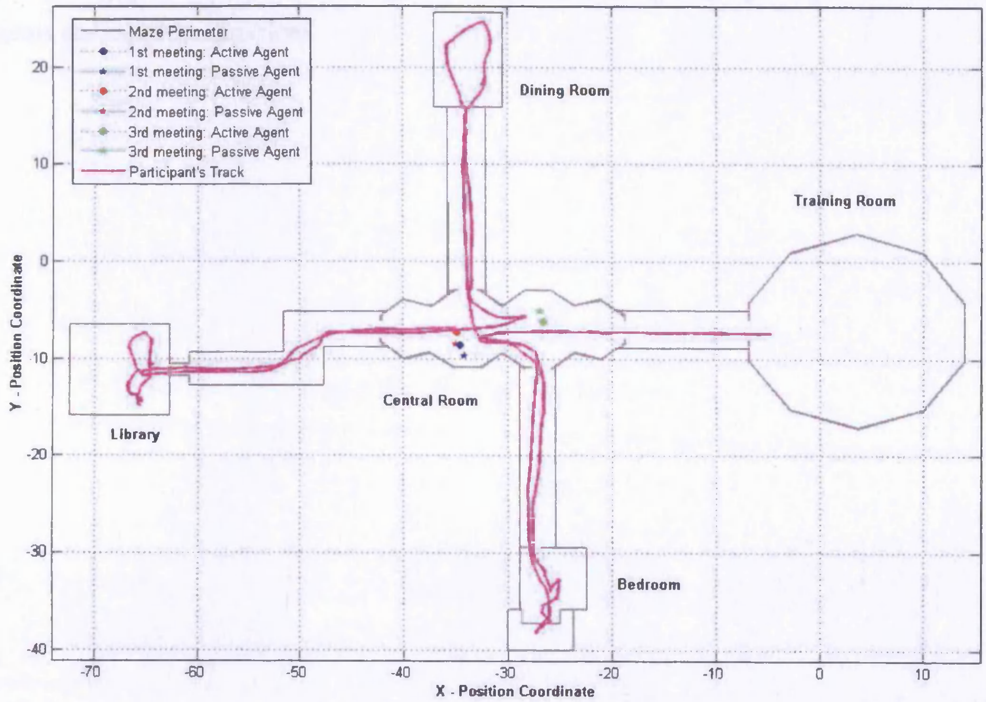


Figure F.18: Participant in the virtual maze

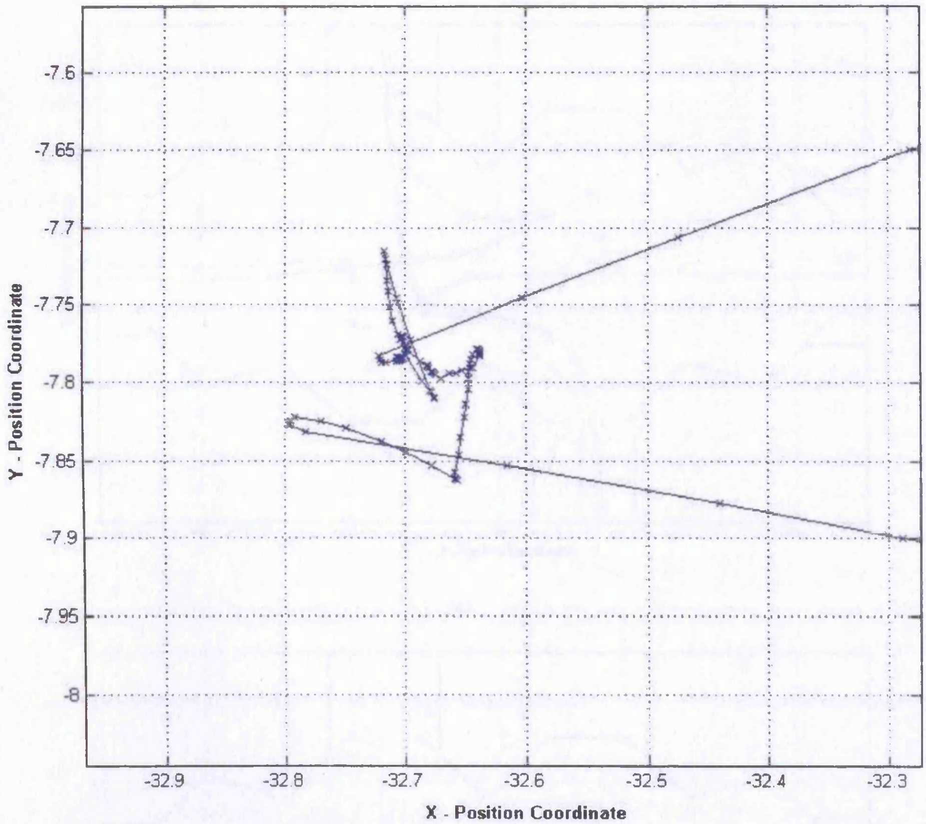
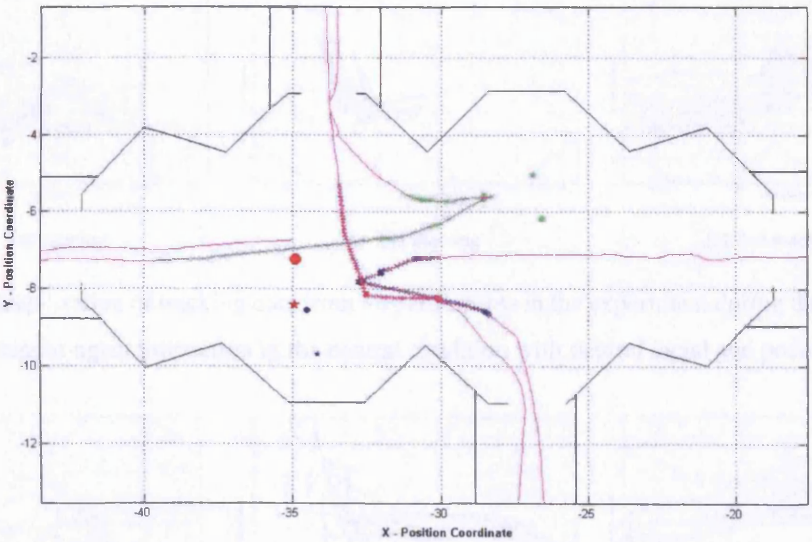
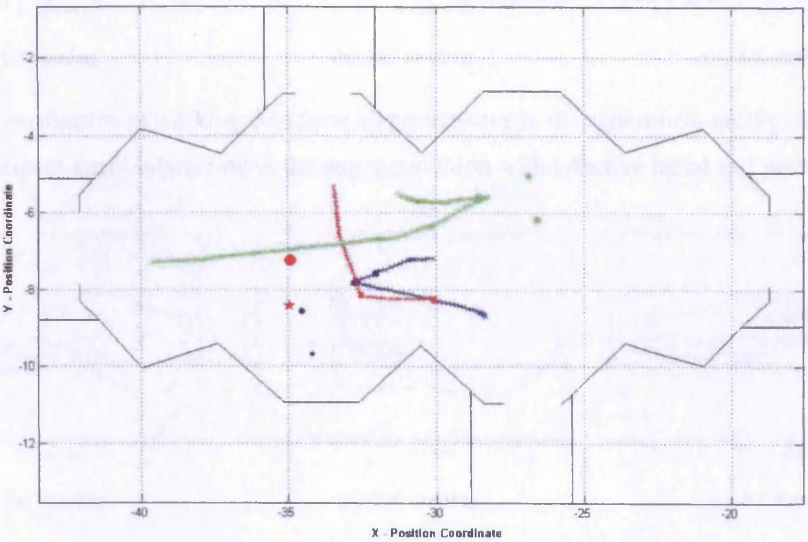


Figure F.19: Sample tracking data showing the characteristic clump where participants stood in front of the agents and asked for directions.

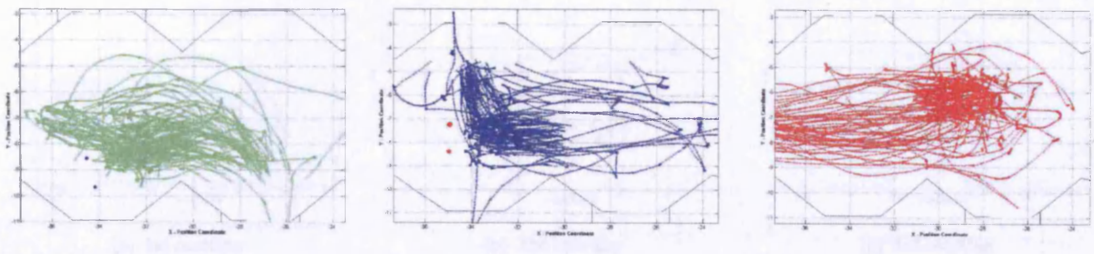


(a)



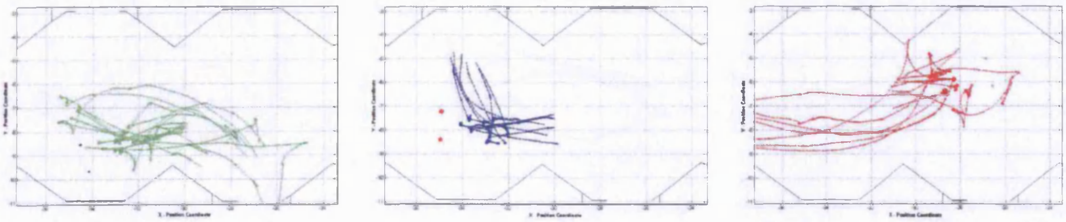
(b)

Figure F.20: Detailed process of extracting tracking data of a participant in the virtual maze



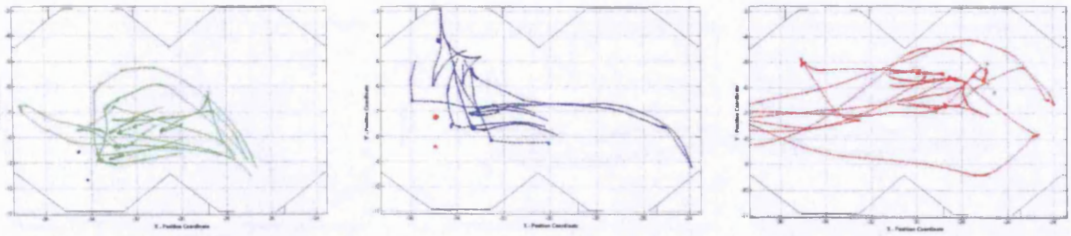
(a) 1st meeting (b) 2nd meeting (c) 3rd meeting

Figure F.21: Visualisation of tracking data from all participants in the experiment during the first, second and third participant-agent interaction.



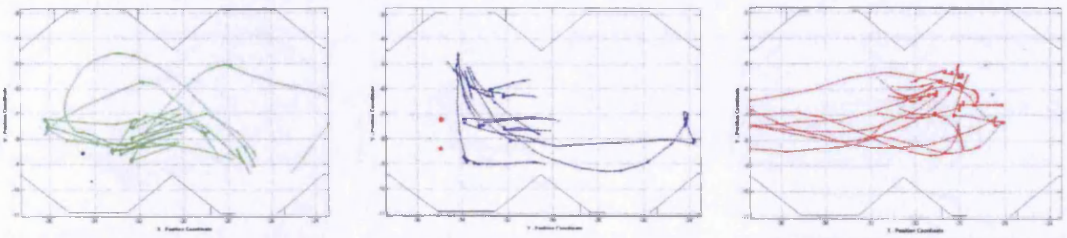
(a) 1st meeting (b) 2nd meeting (c) 3rd meeting

Figure F.22: Visualisation of tracking data from all participants in the experiment during the first, second and third participant-agent interaction in the neutral condition with neutral facial and postural cues.



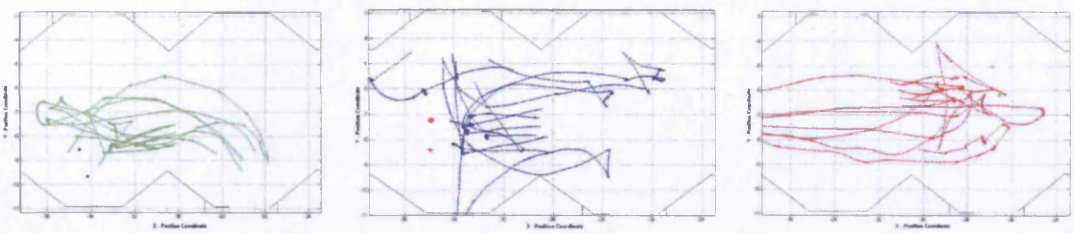
(a) 1st meeting (b) 2nd meeting (c) 3rd meeting

Figure F.23: Visualisation of tracking data from all participants in the experiment during the first, second and third participant-agent interaction in the angry condition with affective facial and postural cues.



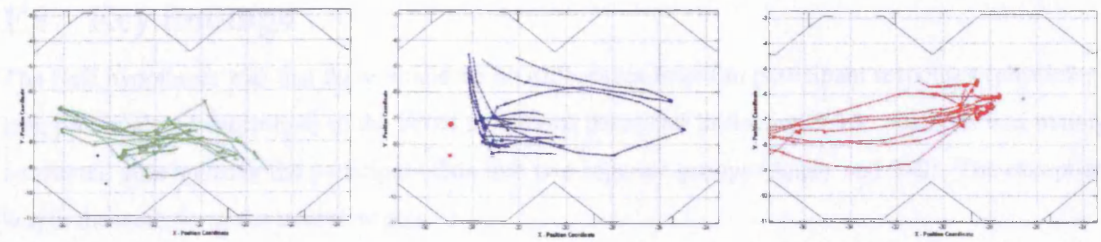
(a) 1st meeting (b) 2nd meeting (c) 3rd meeting

Figure F.24: Visualisation of tracking data from all participants in the experiment during the first, second and third participant-agent interaction in the angry condition with affective facial and neutral postural cues.



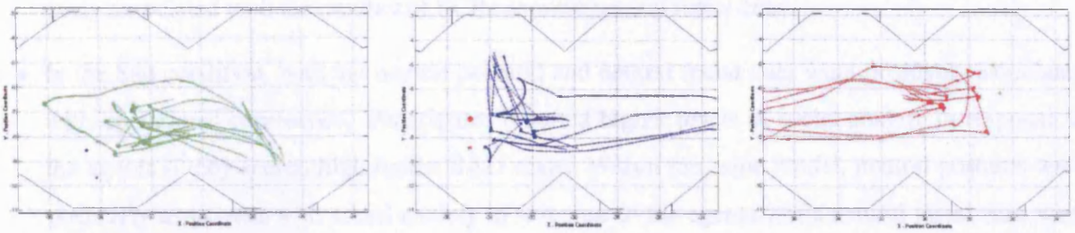
(a) 1st meeting (b) 2nd meeting (c) 3rd meeting

Figure F.25: Visualisation of tracking data from all participants in the experiment during the first, second and third participant-agent interaction in the angry condition with neutral facial and affective postural cues.



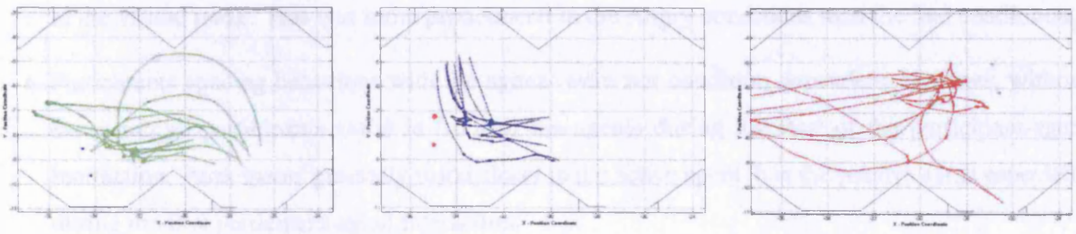
(a) 1st meeting (b) 2nd meeting (c) 3rd meeting

Figure F.26: Visualisation of tracking data from all participants in the experiment during the first, second and third participant-agent interaction in the sad condition with affective facial and postural cues.



(a) 1st meeting (b) 2nd meeting (c) 3rd meeting

Figure F.27: Visualisation of tracking data from all participants in the experiment during the first, second and third participant-agent interaction in the sad condition with affective facial and neutral postural cues.



(a) 1st meeting (b) 2nd meeting (c) 3rd meeting

Figure F.28: Visualisation of tracking data from all participants in the experiment during the first, second and third participant-agent interaction in the sad condition with neutral facial and affective postural cues.

F.8 Key findings

The Null hypothesis was that there would be no differences found in participant responses (physiological, proximal and subjective) to the seven conditions described in Section 5.3.1. Analysis was mainly conducted after splitting the participant data into two separate groups (Angry and Sad). The exception was in the analysis of the interview data.

- In the Angry conditions, the neutral postural cues were associated with increased copresence as reported by participants in the questionnaire based on the operational definition of copresence. Participants accurately recognised the angry active agent's emotional state if the agent displayed congruent behavioural cues (angry postural and angry facial cues). Neutral postures were positively associated with the number of SCRs as were neutral facial cues.

- In the Sad condition, both the neutral postural and neutral facial cues were positively associated with operational copresence. Participants reported higher levels of social anxiety in response to the agents if they scored high on the SAD scale. Within the same model, neutral postures were positively associated with social anxiety in response to the agents while neutral facial cues were negatively associated. Neutral postures were negatively associated with the number of SCRs but neutral facial cues were positively associated with the number of SCRs.

- Spatial distribution analysis of the participants' GSR data showed that independent of condition, participants experienced high physiological arousal while approaching the agents than elsewhere in the virtual maze. This was more pronounced in the Angry conditions than the Sad conditions.

- Participants spacing behaviour with the agents were not condition dependent, however, without exception all participants stood in front of the agents during the start of the participant-agent interaction. Participants generally stood closer to the active agent than the passive agent especially during the first participant-agent interaction.

- Participants' perception of the agents were the result of their interpretation of the scenario and varied quite considerably. This is observable in the manner in which participants judged the emotional states of the agents.

- Most participants reported that the agents were aware of the participants' presence. This perception arose due to the agents' perceived ability to maintain eye contact and also in the agents' perceived knowledge of the virtual maze.

- Most participants reported having realistic responses to the agents such as feeling a sense of intrusion in interpreting the agents and also in feeling the need to be polite to the agents.

- Limited interaction with the agents was stated as the most significant factor leading to reduced copresence. Participants were especially sensitive to mistimed responses.

- A significant number of participants attributed a status of higher authority to the active agent and generally did not pay attention to the passive agent since the passive agent was less responsive.

Appendix G

Experiment on kinesics: Paperwork

G.1 Actor's CV

G.2 Script for the active agent and actor

Every day when we close up the bar, we clean up, cash up and lock up.

We start by collecting all the used glasses. At the same time we pick up all the rubbish, old crisp and cigarette packets. At the same time someone else goes round picking up all the ash trays. The glasses go into the dish washer, while the ash trays get emptied into the bin, rinsed down and then go into the dishwasher.

The biggest job is cleaning the bar. We empty all of the beer overflow trays into the sink then clean the beer pumps. You have to take the pumps apart and scrub the inside bit. You also have to clean out the coffee machine, particularly the tube for foaming the milk as that gets clogged, clean it under running water. You also have to empty all the coffee beans out of the machine as they can't stay in there overnight.

We take all the leftover sliced lemons and oranges, put them in a bowl, cover them with water and then put them in the fridge so that they don't go bad. The lemon and oranges plates, the beer overflow trays, measure glasses, cocktail shakers and other stuff go in the sink and get washed up.

Then we clean out the sink, wipe down the bar and polish the beer pumps. We have to use the soft chamois to polish the beer pumps not the normal cloth. The next thing we have to do is refill the fridges. You have to count out how many bottles of each type are missing and go down to the cellar and pick up replacements.

Once the bar is cleaned, we wipe down all the tables. When they are cleaned we flip the chairs onto the tables and mop the floor. The mop and cloths for cleaning the tables are in the cellar, near the bottom of the stairs. Someone also has to go and check the toilets and wash them up. Make sure you scrub the basins, the mirrors and the toilets seats. You put toilet liquid in all the toilets overnight and then mop up the floors.

After all the cleaning we cash up. That means counting all of the cash in the till and sorting it by note or coin type into bags. You have to make sure that you put the right coins in the right bags, and that there are the right number of coins in each bag, the amount is written on the bag. When it's all counted we put the cash in the cash box and lock it. We hide the key behind the bar in an old cocktail shaker. We then put the cash box into the safe, which is underneath the bar to the left. The combination to the safe is on the inside lid of the cash box. We change it every week. The cash box has to be taken to the bank every morning to deposit the cash.

Once everything else is done we lock up the bar, and pull down the shutters. The shutters are a bit stiff so you have to pull them down really hard.

Every bar manager has a set of keys so they have to do the locking up. We can only go home after doing all those things.

G.3 Possible back stories

G.3.1 Neutral

- You are a barman explaining to a new employee what to do.

G.3.2 Angry

- You are the bar manager talking to a bar man. He's been here for weeks but never does his job properly. What really annoys you is that he doesn't clear up properly as you have to sort it out in the morning. This morning it was really bad, you had to work in a real rush to get everything looking OK before the bar opened, and you had to open half an hour late which meant turning customers away.

- You are a barman. Your patronising boss keeps asking you to repeat over all the things you have to do when closing up the bar. you know this perfectly well, and are sure he's just doing it to piss you off.

- You are a barman explaining to a new employee what to do, except you recognise him as the guy you caught kissing your girlfriend last Friday.

- You are a barman explaining to an employee what to do. He is your younger brother and is known to be irresponsible. Yesterday he took your car without asking and has crashed it.

G.3.3 Sad

- You're a barman; you've been doing this job for years, and have had a heavy week and are sick of it. You are explaining to an employee who is also a mate all the annoying things you have to do every night.

- You are a barman explaining to a new employee what to do. Your mother died yesterday.

- You are a barman explaining to a new employee what to do. He is your older brother. You are going through a very painful and depressing divorce.

G.4 Decoding the participant ID

The first digit gave an indication on which set the participant was assigned to: Angry (1) and Sad (2). The second digit indicated the order in which the two virtual conditions were presented to the participants; 1 indicated that the participant had been exposed to the moving virtual condition followed by the static virtual condition and then the real condition while 2 indicated that the participant had been exposed to the static virtual condition followed by the moving virtual condition and the real condition. The last two digits ranged from '01' to '13' in the case of the Angry set; and from '01' to '10' in the Sad set. The last two digits indicated the batch number.

G.5 Information sheet for Participants

Thank you for participating in our study. This is one of a long series of studies into understanding the responses of people within virtual environments. This study has been approved by *University College London's Committee on the Ethics of Non-NHS Human Research.* Please read through this information sheet and feel free to ask any questions. The experimenters will answer any general questions, however, the specific aspects regarding this study cannot be discussed with you until the end of the session. The whole study will take about *an hour.*

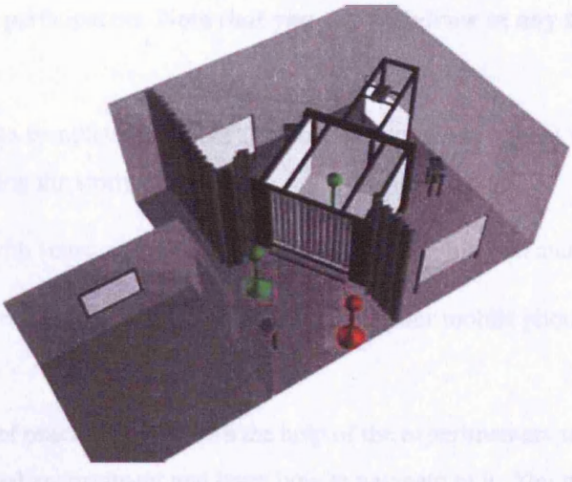


Figure G.1: Model of the VR system

You will be using the CAVE ^{TM}-like system called the ReaCTor. See figure above. The ReaCTor is a VR system made up of 3 walls measuring roughly 3m x 3m x 3m. You will wear VR glasses and be handling a tracking device similar to a joystick. The virtual reality viewing equipment can be worn over eyeglasses. You may be asked to take off your shoes in order to protect the virtual reality equipment. In addition to the tracking equipment used to navigate the system, you will also be fitted with physiological equipment designed to measure your heart rate, respiration and galvanic skin responses whilst you are in the immersive virtual environment.

In this particular study you will be asked to enter a bar **three** times. You will be asked to observe events and form opinions about the people in the bar.

Information that we collect will never be reported in a way that specific individuals can be identified. Information will be reported in a statistical and aggregated manner, and any verbal comments that you make, if written about in subsequent papers, will be presented anonymously.

IMPORTANT

When people use virtual reality systems, some people sometimes experience some degree of nausea. If at any time you wish to stop taking part in the study due to this or any other reason, please just say so and we will stop.

There has been some research, which suggests that people using head-mounted displays might experience some disturbances in vision afterwards. No long-term studies are known to us, but short term studies carried out suggest that even after about 30 minutes of leaving a VR system, some people experience aftermath effects. For this reason, we advise you not to drive or operate heavy machinery for at least three hours of completing the study.

There have been various reported side effects of using virtual reality equipment, such as 'flash-backs'. Also with any type of video equipment there is a possibility that an epileptic episode may be generated. For instance, this has been reported in computer video games.

PROCEDURES

- You will be asked to read, understand and sign a **Consent Form**. If you sign it the study will continue with your participation. **Note that you can withdraw at any time without giving any reasons.**
- You will be asked to complete a number of questions on paper, so that we can try to understand your responses during the study.
- You will be fitted with sensors to measure your heart rate, respiration and galvanic skin responses.
- You will be asked to remove your shoes and switch off your mobile phone(s) before using the VR system.
- You will have a brief practice period with the help of the experimenters so that you can get used to the immersive virtual environment and learn how to navigate in it. You will then go into the maze as mentioned during which you will be videotaped.
- After the visit to the maze you will complete some more questionnaires about your experience.
- Finally there will be a short discussion with the experimenters about your experiences. The interview might be audio or video taped.
- **Thank you** for your participation. Please do not discuss this study with others for about three months, since the study is ongoing.
- Any other questions? *Please **ask** any questions* that come to mind at this point. After this read and sign the **Consent Form**.

In case you have any enquiries regarding this study in the future, please contact:

Vinoba Vinayagamoorthy,
Department of Computer Science,
University College London,

http://www.cs.ucl.ac.uk/staff/V.Vinayagamoorthy

G.6 Consent Form

University College London

Department of Computer Science
Mel Slater
Professor of Virtual Environments

ID

PROJECT	**EQUATOR**
Investigators	Vinoba Vinayagamoorthy, Mel Slater and Anthony Steed
Experiments	Vinoba Vinayagamoorthy

To be completed by volunteers:
We would like you to read the following questions carefully and **circle** your answers.

Have you read the information sheet about this study?	YES/NO
Have you had an opportunity to ask questions and discuss this study?	YES/NO
Have you received satisfactory answers to all your questions?	YES/NO
Have you received enough information about this study?	YES/NO
Which investigator have you spoken to about this study?	…………
Do you understand that you are free to withdraw from this study?	
• At any time	YES/NO
• Without giving a reason for withdrawing	YES/NO
Do you understand and accept the risks associated with the use of virtual reality equipment?	YES/NO
Do you agree to take part in this study?	YES/NO
Do you agree to be video taped?	YES/NO
Do you agree to be audio taped?	YES/NO
Do you agree to be physiological monitored?	YES/NO

I certify that I do **not** have epilepsy.

I certify that I will not be driving a car, motorcycle, bicycle, or use other types of complex machinery that could be a danger to myself or others, within 3 hours after the termination of the study.

Signed…………………………………………………**Date**……………………………….

Name in block letters………………………………………………………………………….

Investigator…………………………………………………………………………………..

In case you have any enquiries regarding this study in the future, please contact:

Vinoba Vinayagamoorthy

Department of Computer Science
University College London

http://www.cs.ucl.ac.uk/staff/V.Vinayagamoorthy

Information that we collect will never be reported in a way that individuals can be identified. Information will be reported in aggregate, and any verbal comments that you make, if written about in subsequent papers, will be presented anonymously.

G.7 Personal information

ID ______________________________

Your Age	
How fluent is your English?	Basic ☐ Proficient ☐ Fluent ☐
Occupational status *(If other, please specify and also your area of interest)*	☐ Undergraduate Student ☐ Masters Student ☐ PhD Student ☐ Research Assistant/Fellow ☐ Staff - systems, technical ☐ Faculty ☐ Administrative Staff ☐ Other
Please state your level of computer literacy on a scale of (1...7) (novice) 1 2 3 4 5 6 7 (expert)	
Please rate your level of experience with computer *programming* (novice) 1 2 3 4 5 6 7 (expert)	
Have you ever experienced 'virtual reality' before? (no experience) 1 2 3 4 5 6 7 (extensive experience)	
How many times did you play video games (at home, work, school, or arcades) in the last year?	Never ☐ 1 - 5 ☐ 6 - 10 ☐ 11 - 15 ☐ 16 - 20 ☐ 21 - 25 ☐ > 25 ☐
How many *hours per week* do you spend playing video games?	0 ☐ < 1 ☐ 1 - 3 ☐ 3 - 5 ☐ 5 - 7 ☐ 7 - 9 ☐ > 9 ☐

G.8 Agent emotion assessment and copresence questionnaire

ID __________ **ORDER** __________

Session [MVP]: Choose **ONE** of the following adjectives to describe each person's attitude towards the **OTHER**

	Surprised	**Afraid**	**Angry**	**Happy**	**Disgusted**	**Sad**	**Neutral**	**Don't know**	**Other**
Right	☐	☐	☐	☐	☐	☐	☐	☐	☐
Left	☐	☐	☐	☐	☐	☐	☐	☐	☐

Other Adjective (R): ________________ Other Adjective (L): ________________

Storyline:

Responses

1. How much did you behave as if the characters were real?

I behaved as if the characters were real...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

2. How much did you find yourself automatically behaving as if the characters were real?

I found myself automatically behaving as if the characters were real...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

3. How much was your emotional response to the characters as if they were real?

My emotional responses to the characters were as if they were real...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

4. How much were your thoughts in relation to the characters as if they were real?

My thoughts of the characters were as if they were real...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

5. How much were you thinking things like 'I know this person isn't real' but then surprisingly finding yourself behaving as if the characters were real?

My behaviour to the characters was as if they were real despite knowing they were not...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

6. How much did you have physical responses (such as change in heart rate, blushing, sweating, etc.) to the characters as if they were real?

My physical responses to the characters were as if they were real ...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

Figure G.2: Questionnaire for the moving virtual condition

Session [SVP]: Choose **ONE** of the following adjectives to describe each person's attitude towards the **OTHER**

	Surprised	**Afraid**	**Angry**	**Happy**	**Disgusted**	**Sad**	**Neutral**	**Don't know**	**Other**
Right	☐	☐	☐	☐	☐	☐	☐	☐	☐
Left	☐	☐	☐	☐	☐	☐	☐	☐	☐

Other Adjective (R): ____________________ Other Adjective (L): ____________________

Storyline:

Responses

1. How much did you behave as if the characters were real?

I behaved as if the characters were real...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

2. How much did you find yourself automatically behaving as if the characters were real?

I found myself automatically behaving as if the characters were real...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

3. How much was your emotional response to the characters as if they were real?

My emotional responses to the characters were as if they were real...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

4. How much were your thoughts in relation to the characters as if they were real?

My thoughts of the characters were as if they were real...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

5. How much were you thinking things like 'I know this person isn't real' but then surprisingly finding yourself behaving as if the characters were real?

My behaviour to the characters was as if they were real despite knowing they were not...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

6. How much did you have physical responses (such as change in heart rate, blushing, sweating, etc.) to the characters as if they were real?

My physical responses to the characters were as if they were real ...

Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

Figure G.3: Questionnaire for the static virtual condition

Session [RP]: Choose **ONE** of the following adjectives to describe each person's attitude towards the **OTHER**

	Surprised	Afraid	Angry	Happy	Disgusted	Sad	Neutral	Don't know	Other
Right	☐	☐	☐	☐	☐	☐	☐	☐	☐
Left	☐	☐	☐	☐	☐	☐	☐	☐	☐

Other Adjective (R): ____________________ Other Adjective (L): ____________________

Storyline:

Based on:

With respect to your responses experienced to the characters, which of the first two sessions was more like the third session with the real person?

FIRST ☐ **SECOND** ☐

These are questions comparing the people in between the three different sessions

How similar was your response to seeing the ***virtual*** person on the ***right*** compared to the ***real*** person on the ***right*** in a similar role? Not at all similar 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very similar
How similar was your response to the ***moving virtual*** person on the ***right*** compared to the ***static virtual*** person on the ***right*** when comparing the first 2 sessions? Not at all similar 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very similar
How similar was your response to the ***real*** person on the ***right*** compared to the ***virtual*** person on the ***right*** in the less similar role? Not at all similar 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very similar
How similar was your response to ***real*** person on the ***right*** looking at you compared to the ***static virtual*** person on the ***right***? Not at all similar 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very similar
How similar was your response to the ***moving virtual*** person on the ***right*** looking at you compared to the ***static virtual*** person on the ***right*** when comparing the first 2 sessions? Not at all similar 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very similar
How similar was your response to ***real*** person on the ***right*** looking at you compared to the ***moving virtual*** person on the ***right***? Not at all similar 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very similar

Figure G.4: Questionnaire for the real condition - Part 1

These are questions comparing the people within the same sessions

How similar was your response to seeing the ***virtual*** person on the ***right*** compared to the ***virtual*** person on the ***left*** in the 1st session? Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much
How similar was your response to seeing the ***virtual*** person on the ***right*** compared to the ***virtual*** person on the ***left*** in the 2nd session? Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much
How similar was your response to seeing the ***real*** person on the ***right*** compared to the ***virtual*** person on the ***left*** in the 3rd session? Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much
How similar was your response to seeing the ***virtual*** person on the ***right*** looking at you compared to the ***virtual*** person on the ***left*** looking at you in the 1st session? Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much
How similar was your response to seeing the ***virtual*** person on the ***right*** looking at you compared to the ***virtual*** person on the ***left*** looking at you in the 2nd session? Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much
How similar was your response to seeing the ***real*** person on the ***right*** looking at you compared to the ***virtual*** person on the ***left*** looking at you in the 3rd session? Not at all 1-☐ 2-☐ 3-☐ 4-☐ 5-☐ 6-☐ 7-☐ Very much

Figure G.5: Questionnaire for the real condition - Part 2

G.9 Pictures

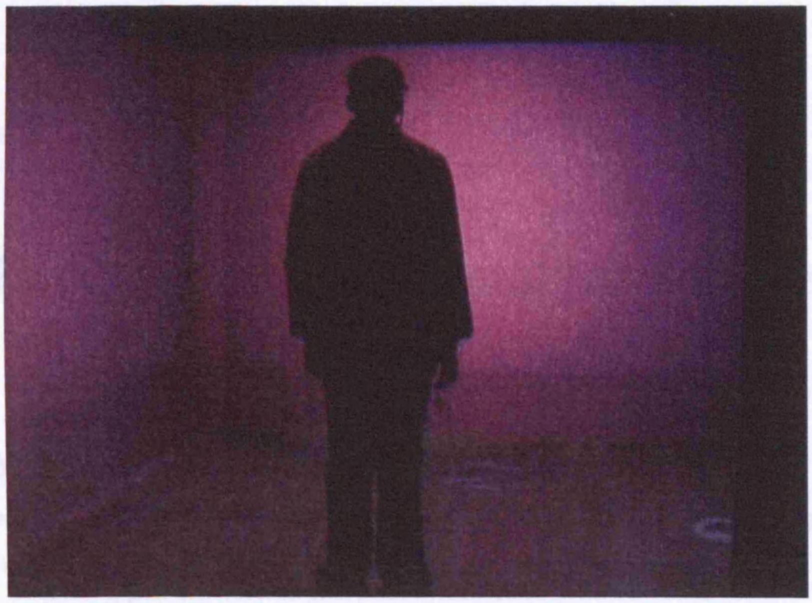


Figure G.6: The training room.

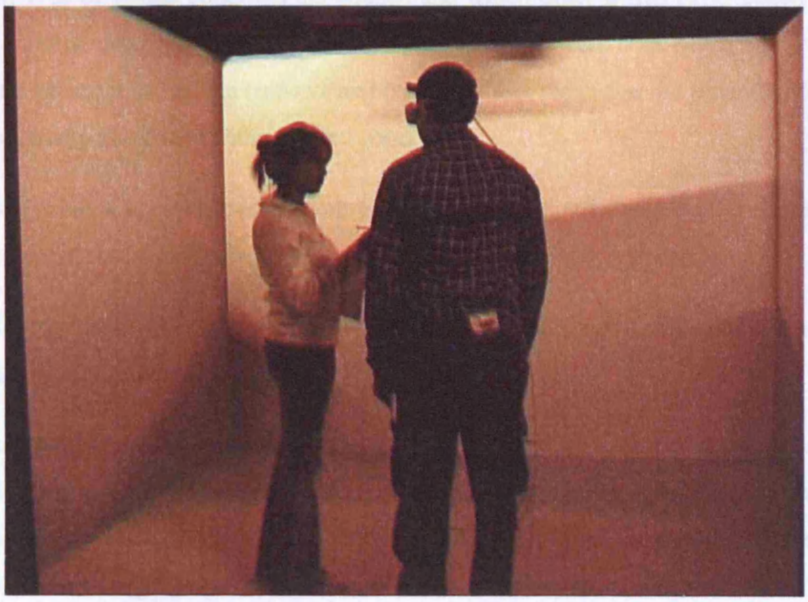


Figure G.7: Questioning and Interviewing a participant in the ReaCTor

Appendix H

Experiment on kinesics: Analysis and Results

H.1 Key findings

The Null hypothesis was that there would be no differences found in participant responses (physiological, proximal and subjective) to the three conditions within each experiment set as described in Section 6.3.1. The key findings from this experiment are listed in this section and discussed in the next section.

- In both the Angry and Sad set, the copresence questionnaire uncovered an interaction effect between the type of virtual condition experience and which order it was experienced in. Although the static virtual condition and the second virtual condition were both negatively associated with reported copresence, when the static virtual condition was viewed as the second virtual condition, it was positively associated with reported copresence.

- In the Angry set, a higher expertise with VR systems led to reduced reported copresence while a higher score on the neuroticism scale was associated with higher copresence.

- In the Sad set, older participants reported higher copresence as did participants who scored high on the extraversion scale.

- In the Angry set, participants reported that their responses in the static virtual condition were least like their responses in the moving virtual condition. Yet the similarity between the moving virtual condition and the real condition was still scored relatively low. Although the score went up slightly if the participants experienced the moving virtual condition first.

- Generally participants were comfortable to subjectively compare and report that the moving virtual condition was in between the static virtual condition and the real condition. However, this was not the case when participant were specifically asked about the moments in which the agents (or actor) turned to look at them. In this case, the participants reported that the two virtual conditions were more alike and very different from the real condition.

- In both the Angry and the Sad sets, the participant's ability to accurately recognise the active agent's emotional state was improved in the moving virtual condition in comparison to the static virtual condition and almost perfect in the real condition.

- In keeping with the experiment reported in chapter 5, participants judged the scenario presented to them holistically. The passive agent's behavioural cues were judge differently in different conditions even though the behaviour of the passive agent never changed.

- The personality measures of the participant played a significant role in improving the explanatory power of all the fitted models presented in this chapter. This was especially the case in the models used to analyse the number of SCRs in the GSR data.

- In both the Angry and the Sad sets, when considering the number of SCRs, there was a strong interaction effect between the type of virtual condition and the order in which the virtual condition was experienced. If the static condition was experienced as the second virtual condition, participants experienced a higher number of SCRs.

- In the Angry sets, the static virtual condition and the second virtual condition were both negatively associated with the number of SCRs. In the Sad set, this effect was not significant.

- In both the Angry and Sad sets, the participant's measure of extraversion and neuroticism were negatively associated with the number of SCRs while their measure of conscientiousness was positively associated with the number of SCRs.

- In the Angry set, the participant's measure of openness to new experiences was positively associated with the number of SCRs but in the Sad set, the inverse association was true.

- Additionally, in the Angry set, the participant's measure of agreeableness was negatively associated with the number of SCRs.

- In the event-related analysis, visual analysis of the GSR data around the moments when the agents (or actor) looks at the participant revealed a wide variety of possible responses. The actor elicited a much greater response from the participant followed by the active agent in the moving virtual condition. However, the actor clearly outperformed the active agent.

Appendix I

Miscellaneous

I.1 Properties of agents

Property	Meaning
character	possessing a plausible personality and emotional state
reactive	responds in a timely fashion to changes in the environment
autonomous	exercises control over its own actions
communicative	socially able, communicates with other agents and participants
goal-oriented	purposeful, does not simply act in response to the environment
learning	adaptive changes its behaviour based on its previous experience
flexible	actions are not scripted

Table I.1: Properties of agents as defined by researchers in the field

I.2 Hall's proxemic model

Intimate distance: Situations in the ***close phase*** require physical contact, high involvement and awareness e.g. comforting and protective behaviours with most vocalisation being involuntary. Situations in the ***far phase*** are usually uncomfortable and crowded with low-level whispers.

Personal distance: Interactions in the ***close phase*** are normally used to signal a close relationship. The saying '*keeping one at arm's length*' occurs within the ***far phase*** of this zone signalling the limit of *physical* domination. The voice level used is moderate and subject matters discussed in the far phase are of personal interests and normally requires involvement.

Social distance: The ***close phase*** in this space is the limit for psychological domination. Impersonal businesses and socialisation at gatherings are conducted within this phase. The effect obtained by standing in an individual's social ***close*** distance and looking at them is domineering. Situations involving business and social discourse are conducted within the ***far phase*** in the social space. Vocalisations are loud.

Public distance: The ***close phase*** within this space is the distance at which an individual takes evasive action in the face of a threat. It is a subliminal form of the '*fight or flight*' phenomena. Vocalisations are loud but not at full volume. In the ***far phase*** of this zone, most subtle shades in verbal communication and facial expressions are lost so nonverbal communication almost entirely shifts to postural cues and gestures. Vocalisations are well enunciated.

I.3 The ideal human model

These measurements are adapted from Larmann (2006). They give some standards for measurement of the human body which have been established by artists working with human models.

I.3.1 Standard Bodily Proportions

The human figure is an average of 7 heads high. The width from shoulder to shoulder is 3 heads width. The distance from the hip to the toes is 4 heads. The distance from the top of the head to the bottom of the chest is 2 heads. The distance from the wrist to the end of the outstretched fingers of the hand is 1 head. The length from top to bottom of the buttocks is 1 head. The distance from the elbow to the end of outstretched fingers is 2 heads.

I.3.2 Standard Facial Proportions

The eyes are halfway between the top of the head and the chin. The bottom of the nose is halfway between the eyes and the chin. The mouth is halfway between the nose and the chin. The corners of the mouth line up with the centres of the eyes. The top of the ears line up above the eyes, on the eyebrows. The bottom of the ears line up with the bottom of the nose.

Appendix J

List of Acronyms

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ANS	Autonomic Nervous System
AU	Action Unit
BVH	Biovision Hierarchical
BVP	Blood Volume Pulse
DIVE	Distributed Interactive Virtual Environment
ECG	Electrocardiogram
EDA	Electrodermal Activity
EEG	Electroencephalogram
EMG	Electromyogram
EOG	Electrooculography
FFM	Five-Factor Model
GSR	Galvanic Skin Response
GUI	Graphical User Interface
HMD	Head Mounted Display
HRV	Heart Rate Variability
IVE	Immersive Virtual Environment
OCC	Ortony, Clore and Collins
PAD	Pleasure-Arousal-Dominance
PNS	Parasympathetic Nervous System
PIAVCA	Platform Independent Architecture for Virtual Characters and Avatars
SNS	Sympathetic Nervous System
VE	Virtual Environment
VGA	Video Graphics Array
VR	Virtual Reality
VRPN	Virtual-Reality Peripheral Network

Bibliography

Albrecht, I., Schröder, M., Haber, J., and Seidel, H.-P. (2005). Mixed feelings: Expression of non-basic emotions in a muscle-based talking head. *Special issue of Journal of Virtual Reality on Language, Speech & Gesture*, 8:201–212.

Amaya, K., Bruderlin, A., and Calvert, T. (1996). Emotion from motion. In *Graphics Interface*, pages 222–229. Canadian Information Processing Society / Canadian Human-Computer Communications Society.

André, E., Klesen, M., Gebhard, P., Allen, S., and Rist, T. (1999). Integrating models of personality and emotions into lifelike characters. In *Proceedings of the Workshop on Affect in Interactions Towards a new Generation of Interfaces*, pages 136–149.

Andreassi, J. L. (1995). *Psychophysiology: human behavior and physiological response*. Lawrence Erlbaum associates.

Argyle, M. (1969). *Social interaction*. Tavistock publications.

Argyle, M. (1998). *Bodily Communication*. Methuen & Co Ltd, second edition.

Argyle, M. and Cook, M. (1976). *Gaze and mutual gaze*. Cambridge University Press.

Argyle, M. and Dean, J. (1965). Eye contact, distance and affiliation. *Sociometry*, 28:289–304.

Argyle, M. and Ingham, R. (1972). Mutual gaze and proximity. *Semiotica*, 6:32–49.

Argyle, M., Ingham, R., Alkema, F., and McCallin, M. (1973). The different functions of gaze. *Semiotica*, 7:10–32.

Argyle, M., Lefebvre, L., and Cook, M. (1974). The meaning of five patterns of gaze. *European Journal of Social Psychology*, 4(2):125–136.

Argyle, M. and Trower, P. (1980). *Person to person: ways of communicating*. Harper and Row.

Arts, E. (2006). The sims online, http://thesims.ea.com/.

ATLAS.ti (2006). Atlas.ti - the knowledge workbench, http://www.atlasti.com/index.php.

Autodesk (2006). Autodesk - autodesk 3ds max - overview, http://www4.discreet.com/3dsmax/.

Badler, N., Allbeck, J., Zhao, L., and Byun, M. (2002). Representing and parameterizing agent behaviors. In *Proceeding Computer Animation*, pages 133–143, Geneva, Switzerland. IEEE Computer Society June 2002.

Badler, N. I., Hollick, M. J., and Granieri, J. P. (1993). Real-time control of a virtual human using minimal sensors. *Presence: Teleoperators and Virtual Environments*, 2(1):82–86.

Bahill, T. A., Adler, D., and Stark, L. (1975). Most naturally occurring human saccades have magnitudes of 15 degrees or less. *Investigative Ophthalmology*, 14:468–469.

Bailenson, J. N. and Blascovich, J. (2004). *Avatars*, pages 64–68. Encyclopedia of Human-Computer Interaction. Berkshire Publishing Group, first edition.

Bailenson, J. N., Blascovich, J., and Beall, A. C. (2001). Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators and Virtual Environments*, 10(6):583–598.

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, 29(7):1–15.

Bartneck, C. (2002). Integrating the occ model of emotions in embodied characters. In *Proceedings of the Workshop on Virtual Conversational Characters*, pages 39–48.

Bartneck, C. and Reichenbach, J. (2005). Subtle emotional expressions of synthetic characters. *International journal of human-computer studies*, 62(2):179–192.

Bates, J., Loyall, B. A., and Reilly, S. N. (1994). An architecture for action, emotion and social behaviour. In *Selected papers from the 4th European Workshop on on Modelling Autonomous Agents in a Multi-Agent World, Artificial Social Systems*, pages 55–68. Springer-Verlag.

Bécheiraz, P. and Thalmann, D. (1996). A model of nonverbal communication and interpersonal relationship between virtual actors. *Computer Animation*, pages 58–67.

Becker, C., Kopp, S., and Wachsmuth, I. (2004). Simulating the emotion dynamics of a multimodal conversational agent. In *Affective Dialogue Systems*, pages 154–165.

Benford, S., Bowers, J., Fahlen, L. E., Greenhalgh, C., and Snowdon, D. (1995). User embodiment in collaborative virtual environments. In *Proceedings of ACM SIGCHI*, pages 242–249.

Bickmore, T. and Picard, R. (2006). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer Human Interaction (ToCHI)*, 12(2):293–327.

Birdwhistell, R. L. (1971). *Kinesics and context: essays on body motion communication*. Allen Lane The Penguin Press.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K., Hoyt, C., and Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychology Inquiry*, 13:103–124.

Brand, M. and Hertzmann, A. (2000). Style machines. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques (SIGGRAPH)*, pages 183–192.

Breese, J. S. and Ball, G. (1998). Modeling emotional state and personality for conversational agents. Technical Report MSR-TR-98-41.

Brewer, M. B. and Hewstone, M. (2004). *Emotion and motivation*. Blackwell Publishing.

Brogni, A., Slater, M., and Steed, A. (2003). More breaks less presence. In *The 6th Annual International Workshop on Presence*.

Brogni, A., Vinayagamoorthy, V., Steed, A., and Slater, M. (2006). Variations in physiological responses of participants during different stages of an immersive virtual environment experiment. In *Proceedings of the ACM symposium on Virtual Reality Software and Technology (VRST)*.

Brosschot, J. F. and Thayer, J. F. (2003). Heart rate response is longer after negative emotions than after positive emotions. *International Journal of Psychophysiology*, 50(3):181–187.

Bruderlin, A. and Williams, L. (1995). Motion signal processing. *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques (SIGGRAPH)*, pages 97–104.

Brunner, L. J. (1979). Smiles can be back channels. *Journal of Personality and Social Psychology*, 37(5):728–734.

Bull, P. E. (1987). *Posture and Gesture*, volume 16 of *International Series in Experimental Social Psychology*. Pergamon Press.

Cahn, J. E. (1990). Generating expression in synthesised speech. Technical report.

Carroll, J. M. and Russell, J. A. (1996). Do facial expressions signal specific emotions? judging emotion from the face in context. *Journal of personality and social psychology*, 70(2):205–218.

Cassell, J., Bickmore, T. W., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H. H., and Yan, H. (1999). Embodiment in conversational interfaces: Rea. In *Proceedings of ACM SIGCHI*, pages 520–527. ACM Press.

Cassell, J., Nakano, Y. I., Bickmore, T. W., Sidner, C. L., and Rich, C. (2001). Non-verbal cues for discourse structure. In *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics*, pages 114–123. Association for Computational Linguistics.

Cassell, J., Pelachaud, C., Badler, N. I., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S., and Stone, M. (1994). Animated conversation: Rule-based generation of facial expression, gesture

& spoken intonation for multiple conversational agents. In *Proceedings of ACM SIGGRAPH*, pages 413–420.

Cavazza, M., Charles, F., and Mead, S. J. (2002). Character-based interactive storytelling. In *IEEE Intelligent systems*, pages 17–24.

Chi, D., Costa, M., Zhao, L., and Badler, N. I. (2000). The emote model for effort and shape. In *Proceedings of ACM SIGGRAPH*, volume ACM Computer Graphics Annual Conference, pages 173–182.

Chittaro, L. and Serra, M. (2004). Behavioral programming of autonomous characters based on probabilistic automata and personality. *Journal of Visualization and Computer Animation*, 15(3-4):319–326.

Colburn, A. R., Cohen, M. F., and Drucker, S. M. (2000). The role of eye gaze in avatar mediated conversational interfaces. Technical Report MSR-TR-2000-81.

Collet, C., Vernet-Maury, E., Delhomme, G., and Dittmar, A. (1997). Autonomic nervous system response patterns specificity to basic emotions. *Journal of autonomic nervous system*, 62(1-2):45–57.

Cosatto, E. and Graf, H. P. (2000). Photo-realistic talking heads from image samples. In *IEEE transactions on Multimedia*, volume 2, pages 152–163.

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, 28(2):117–139.

Creswell, J. W. (2003). *Research design: qualitative, quantitative and mixed methods approaches*. Sage publications, second edition.

Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: The design and implementation of the cave. In *Proceedings of ACM SIGGRAPH*, pages 135–142.

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals (Orignially published in 1872, reprinted in 1999)*. Fontana Press, third edition.

Davies, E. (2001). *Beyond Dance, Laban's Legacy of Movement Analysis*. Berchin Books.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7:242–249.

De Silva, R. P., Kleinsmith, A., and Bianchi-Berthouze, N. (2005). Towards unsupervised detection of affective state nuances from body postures. In Tao, J., Tan, T., and Picard, R. W., editors, *Proceedings of the First International on Conference on Affective Computing & Intelligent Interaction*, volume 3784 of *Lecture Notes in Computer Science*, pages 32–39. Springer.

Deng, Z., Lewis, J. P., and Neumann, U. (2003). Practical eye movement model using texture synthesis. In *In Proceedings of ACM SIGGRAPH: Sketches and Applications*.

Deng, Z., Lewis, J. P., and Neumann, U. (2005). Automated eye motion using texture synthesis. *IEEE Computer Graphics and Applications*, 25(2):24–30.

Disney (1937). Snow white and the seven dwarfs. 16mm 83 min, Walt Disney.

Dittmann, A. T., Parloff, M. B., and Boomer, D. S. (1965). Facial and bodily expression: A study of receptivity of emotional cues. *Psychiatry*, 28:239–244.

Doherty, W. (1997). The emotional contagion scale: A measure of individual differences. *Journal of Nonverbal Behaviour*, 21(2):131–154.

Duck, S. (1998). *Human relationships*. SAGE, third edition.

Duclos, S. E., Laird, J. D., Schneider, E., Sexter, M., Stern, L., and Van Lighten, O. (1989). Emotion-specific effects of facial expressions and postures on emotional experience. *Journal of Personality and Social Psychology*, 57(1):100–108.

Edlinger, G., Guger, C., Leeb, R., Friedman, D., Vinayagamoorthy, V., and Slater, M. (2006). Controlling virtual environments by thoughts. In *The 5th Forum of European Neurosciences (Poster presentation in the Cortex and Thalamus session)*.

Egges, A., Kshirsagar, S., and Magnenat-Thalmann, N. (2002). Imparting individuality to virtual humans. In *First International Workshop on Virtual Reality Rehabilitation (Mental Health, Neurological, Physical, Vocational)*.

Egges, A., Kshirsagar, S., and Magnenat-Thalmann, N. (2004). Generic personality and emotion simulation for conversational agents. *Computer Animation and Virtual Worlds*, 15(1):1–13.

Egges, A. and Magnenat-Thalmann, N. (2005). Emotional communicative body animation for multiple characters. In *Proceedings of the First International Workshop on Crowd Simulation*, pages 31–40.

Ekman, P. (1965). *Communication through nonverbal behaviour: A source of information about an interpersonal relationship*, chapter XIII, pages 390–442. Affect, cognition and personality: empirical studies eds. S. S. Tomkins and C. E. Izard. Springer publishing company, inc.

Ekman, P. (1982). *Emotion in the Human Face*, volume Second. Cambridge University Press, New York.

Ekman, P. (1992). Facial expressions of emotion: new findings, new questions. *Psychological science*, 3(1):34–38.

Ekman, P. and Davidson, R. J. (1994). *The nature of emotion: fundamental questions*. Series in affective science. Oxford Unviersity Press.

Ekman, P. and Friesen, V. W. (1967). Head and body cues in the judgement of emotion: A reformulation. *Perceptual and Motor Skills*, 24:711–724.

Ekman, P. and Friesen, W. V. (1969). The repertoire of nonverbal behaviour: categories, origins, usage and coding. *Semiotica*, 1:49–98.

Ekman, P. and Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial cues*. Prentice-Hall.

Ekman, P. and Friesen, W. V. (1976). *Pictures of facial affects*. Consulting psychologist press, Palo Alto.

Ekman, P. and Friesen, W. V. (1978). *Facial Action Coding System*. Consulting Psychologist Press, Inc.

Ekman, P., Levenson, R. W., and Friesen, W. V. (1983). Autonomic nervous system activity distinguishes between emotions. *Science*, 221(4616):1208–1210.

El-Nasr, M. S., Ioerger, T. R., and Yen, J. (1999a). Peteei: a pet with evolving emotional intelligence. In *Proceedings of the third annual conference on Autonomous Agents*, pages 9–15. ACM Press.

El-Nasr, M. S., Ioerger, T. R., and Yen, J. (1999b). A web of emotions. In *Proceedings of the Workshop on Emotion based Agent Architecture (in conjunction with Autonomous Agents)*.

El-Nasr, M. S., Yen, J., and Ioerger, T. R. (2000). Flame - fuzzy logic adaptive model of emotions. In *Autonomous Agents and Multi-Agent Systems*, volume 3, pages 219–257. Kluwer Academic Publishers.

Ellsworth, P. C. and Carlsmith, M. J. (1971). *Effects of eye contact and verbal content on affective respons to a dyadic interaction.*, chapter 6, pages 64–71. Readings in Experimental Social Psychology. Meredith Corporation.

Exline, R. V. and Winters, L. C. (1965). *Affective relations and mutual glances in dyads*, chapter XI, pages 319–350. Affect, cognition and personality: empirical studies eds. S. S. Tomkins and C. E. Izard. Springer publishing company, inc.

Fehr, B. and Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, 113(3):464–486.

Feldman, P. J., Cohen, S., Hamrick, N., and Lepore, S. J. (2004). Psychological stress, appraisal, emotion and cardiovascular response in a public speaking task. *Psychology and health*, 19(3):353–368.

Flack Jr, W. F., Laird, J. D., and Cavallaro, L. A. (1999). Separate and combined effects of facial expressions and bodily postures on emotional feelings. *European Journal of Social Psychology*, 29(2):203–217.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., and Dutton, K. (2000). Facial expressions of emotion: are angry faces detected more efficiently. *Cognition and emotion*, 14(1):61–92.

Frécon, E., Smith, G., Steed, A., Stenius, M., and Stahl, O. (2001). An overview of the coven platform. *Presence: Teleoperators and Virtual Environments*, 10(1):109–127.

Freeman, D., Slater, M., Bebbington, P., Garety, P. A., Kuipers, E., Fowler, D., Met, A., Read, C. M., Jordan, J., and Vinayagamoorthy, V. (2003). Can virtual reality be used to investigate persecutory ideation? *Journal of Nervous and Mental Disease*, 191(8):509–514.

Freeman, J., Avons, S. E., Meddis, R., Pearson, D. E., and IJsselsteijn, W. (2000). Using behavioral realism to estimate presence: A study of the utility of postural responses to motion-stimuli. *Presence: Teleoperators and Virtual Environments*, 9(2):149–164.

Freeman, J., Avons, S. E., Pearson, D. E., and IJsselsteijn, W. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoperators and Virtual Environments*, 8(1):1–13.

Frijda, N. H. (1986). *The emotions*. Cambridge university press.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, 43(5):349–358.

Fukayama, A., Ohno, T., Mukawa, N., Sawaki, M., and Hagita, N. (2002). Messages embedded in gaze of interface agents — impression management with agent's gaze. In *Proceedings of ACM SIGCHI*, pages 41–48.

Fukayama, A., Sawaki, M., Ohno, T., Murase, H., Hagita, N., and Mukawa, N. (2001). Expressing personality of interface agents by gaze. In *Proceedings of INTERACT*, pages 793–794.

Garau, M. (2003). *The Impact of Avatar Fidelity on Social Interaction in Virtual Environments*. PhD thesis.

Garau, M., Slater, M., Bee, S., and Sasse, M. A. (2001). The impact of eye gaze on communication using humanoid avatars. In *Proceedings of ACM SIGCHI*, pages 309–316.

Garau, M., Slater, M., Pertaub, D.-P., and Razzaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 14(1):104–116.

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., and Sasse, A. M. (2003). The impact of avatar realism and eye gaze control on the perceived quality of communication in a shared immersive virtual environment. In *Proceedings of ACM SIGCHI*, pages 529–536.

Garau, M., Widenfeld, H. R., Antley, A., Friedman, D., Brogni, A., and Slater, M. (2004). Temporal and spatial variations in presence: A qualitative analysis. In *The 7th Annual International Presence Workshop*, pages 232–239.

Garcia-Rojas, A., Gutierrez, M., Thalmann, D., and Vexo, F. (2005). Multimodal authoring tool for populating a database of emotional reactive animations. In *Proceedings of the 2nd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms*.

Gard, K. A., Gard, G. C., Dossett, D., and Turone, R. (1982). Accuracy in nonverbal communication as affected by trait and state anxiety. *Perceptual and Motor Skills*, 55(3 pt. 1):747–753.

Gebhard, P. (2005). Alma: a layered model of affect. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 29–36. ACM Press.

Gilbert, T. and Coulson, M. (2002). The role of emotional perception in attentional bias for threat. In *Proceedings of the British Psychological Society*, volume 10, page 68.

Gillies, M. and Ballin, D. (2003). A model of interpersonal attitude and posture generation. In Rist, T., Aylett, R., Ballin, D., and Rickel, J., editors, *Proceedings of the Fourth Workshop on Intelligent Virtual Agents (Lecture Notes in Computer Science)*, volume 2792 of *Lecture Notes in Computer Science*, pages 88–92. Springer.

Gillies, M. and Ballin, D. (2004). Integrating autonomous behaviour and user control for believable agents. In *Proceedings of the Third international joint conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 336–343. IEEE Computer Society.

Gillies, M. and Slater, M. (2005). Non-verbal communication for correlational characters. In Slater, M., editor, *Proceedings of the 8th Annual International Workshop on Presence*, pages 103–106.

Gillies, M., Vinayagamoorthy, V., Robertson, D., and Steed, A. (2005). A platform independent architecture for virtual characters and avatars. In Panayiotopoulos, T., Gratch, J., Aylett, R., Ballin, D., Olivier, P., and Rist, T., editors, *Proceedings of the Third International Workshop on Intelligent Virtual Agents (Poster presentation: Lecture Notes in Computer Science)*, volume 3661, pages 492–492. Springer-Verlag.

Glassner, B. and Moreno, J. D. (1989). *The qualitative-quantitative distinction in the social sciences*, volume 112 of *Boston studies in the philosophy of science*. Kluwer academic publishers.

Goffman, E. (1963). *Behavior in public places: notes on the social organization of gatherings*. Free Press.

Gratch, J. (2000). Émile: Marshalling passions in training and education. In *Proceedings of the 4th International conference of Autonomous Agents*, pages 325–332.

Gratch, J. and Marsella, S. (2001). Tears and fears: modeling emotions and emotional behaviors in synthetic agents. In *Proceedings of the fifth international conference on Autonomous agents*, pages 278–285. ACM Press.

Gratch, J. and Marsella, S. (2004). A domain independent framework for modeling emotion. *Journal of Cognitive Systems Research*, 5(4):269–306.

Gratch, J. and Marsella, S. (2005). Lessons from emotion psychology for the design of lifelike characters. *Applied Artificial Intelligence (special issue on Educational Agents - Beyond Virtual Tutors)*, 19(3-4):215–233.

Gratch, J., Rickel, J., André, E., Cassell, J., Petajan, E., and Badler, N. I. (2002). Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems*, 17(4):54–63.

Green, M. J., Williams, L. M., and Davidson, D. (2003). In the face of danger: specific viewing strategies for facial expressions of threat? *Cognition and emotion*, 17(5):779–786.

Greene, J. C., Caracelli, V. J., and Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3):255–274.

Gregory, W. (2006). William gregory, http://www.williamgregory.co.uk/.

Guger, C., Edlinger, G., Leeb, R., G, P., Antley, A., Garau, M., Brogni, A., Friedman, D., and Slater, M. (2004). Heart-rate variability and event-related ecg in virtual environments presence. In *Presence 2004*, volume The 7th Annual International Workshop on Presence, pages 240–245.

Guye-Vuillème, A., Capin, K. T., Pandzic, S. I., Magnenat-Thalmann, N., and Thalmann, D. (1999). Nonverbal communication interface for collaborative virtual environments. *The Virtual Reality Journal*, 4:49–59.

H-Anim (2006). The humanoid animation working group home page, http://www.hanim.org/.

Hall, E. T. (1969). *The hidden dimension: man's use of space in public and private.* Bodley Head.

Hall, L. and Woods, S. (2005). *Empathic interaction with synthetic characters: the importance of similarity.* Encyclopedia of Human-Computer Interaction. Idea Group Publishing.

Hall, L., Woods, S., Aylett, R., Newall, L., and Paiva, A. (2005). Achieving empathic engagement through affective interaction with synthetic characters. In Tao, J., Tan, T., and Picard, R. W., editors, *Proceedings of the First International Conference on Affective Computing and Intelligent Interaction*, volume 3784, pages 731–738. Springer-Verlag.

Hansen, C. H. and Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. *Journal of personality and social psychology*, 54(6):917–924.

Hatfield, E., Cacioppo, J. J., and Rapson, R. (1994). *Emotional contagion.* Cambridge Unviersity Press.

Held, R. M. and Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators and Virtual Environments*, 1(1):109–112.

Herbelin, B., Benzaki, P., Riquier, F., Renault, O., and Thalmann, D. (2004). Using physiological measures for emotional assessment: a computer-aided tool for cognitive and behavioural therapy. In *Proceedings of ICDVRAT 2004, 5th International Conf on Disability, Virtual Reality*, pages 307–314.

Hess, U., Blairy, S., and Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behaviour*, 21(4):241–257.

Hewes, G. (1957). The anthropology of posture. *Scientific American*, 196:123–132.

Hsu, E., Pulli, K., and Popović, J. (2005). Style translation for human motion. *Proceedings of ACM SIGGRAPH*, 24(3):1082–1089.

Isbister, K. and Nass, C. (2000). Consistency of personality in interactive characteristics: verbal cues, non-verbal cues and user characteristics. *International Journal of Human-Computer Studies*, 53:251–267.

Isen, A. M. (1987). *Positive affect, cognitive processes and social behaviour*, volume 20 of *Advances in experimental social psychology*, pages 203–253. Academic Press.

Isen, A. M., Daubman, K. A., and Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of personality and social psychology*, 52(6):1122–1131.

Izard, C. E. (1971). *The face of emotion*. Appleton-Century-Crofts.

Izard, C. E. (1992). Basic emotions, relations among emotions and emotion-cognition relations. *Psychological review*, 99(3):561–565.

James, W. T. (1932). A study of the expression of bodily posture. *Journal of general psychology*, 42:90–94.

Jang, D. P., Kim, I. Y., Nam, S. W., Wiederhold, B. K., Wiederhold, M. D., and Kim, S. I. (2002). Analysis of physiological response to two virtual environments: Driving and flying simulation. In *CyberPsychology & Behaviour*, volume 5, pages 11–18.

Jeffrey, P. (1998). Personal space in a virutal community (student poster). In *Proceedings of ACM SIGCHI*, pages 347–348.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and psychophysics*, 14:201–211.

Johnsen, K., Dickerson, R., Raij, A., Lok, B., Jackson, J., Shin, M., Hernandez, J., Stevens, A., and Lind, D. S. (2005). Experiences in using immersive virtual characters to educate medical communication skills. In *Proceedings of the IEEE Virtual Reality Conference*, volume 324, pages 179–186.

Johnson, J. A. (2006). Personality test center - ipip neo five factor model of personality, http://www.personalitytest.net/ipip/ipipneo120.htm.

Johnson, M. P. (2003). *Exploiting Quaternions to Support Expressive Interactive Character Motion*. PhD thesis.

Keltner, D. and Buswell, B. N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition and Emotion*, 10(2):155–171.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:22–63.

Kendon, A. (1970). Movement coordination in social interaction: some examples described. *Acta Psychologica*, 32(2):1–25.

Kesteren, A.-J. v., Akker, R. o. d., Poel, M., and Nijholt, A. (2000). Simulation of emotions of agents in virtual environments using neural networks. In *Learning to Behave: Internalising Knowledge. Twente Workshop on Language Technology 18*, pages 137–147.

Kleinginna, P. and Kleinginna, A. (1981). A categorized list of emotion definitions with suggestions for a consensual definition. *Motivation and Emotion*, 5:345–379.

Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Presence: Teleoperators and Virtual Environments*, 100(1):78–100.

Kleinsmith, A., De Silva, R. P., and Bianchi-Berthouze, N. (2005). Grounding affective dimensions into posture features. In Tao, J., Tan, T., and Picard, R. W., editors, *Proceedings of the First International on Conference on Affective Computing & Intelligent Interaction*, volume 3784 of *Lecture Notes in Computer Science*, pages 263–270. Springer.

Kshirsagar, S. and Magnenat-Thalmann, N. (2002a). A multilayer personality model. In *Proceedings of the 2nd International Symposium on Smart Graphics*, pages 107–115.

Kshirsagar, S. and Magnenat-Thalmann, N. (2002b). Virtual humans personified. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 356–357. ACM Press.

Kurlander, D., Skelly, T., and Salesin, D. (1996). Comic chat. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (SIGGRAPH)*, pages 225–236. ACM Press.

Langton, S. R. H., Watt, R. J., and Bruce, V. (2000). Do the eyes have it? cues to the direction of social attention. *Trends in Cognitive Sciences*, 4(2):50–59.

Larmann, R. (2006). The figure drawing laboratory, university of evansville, http://www2.evansville.edu/drawinglab.

Lasseter, J. (1987). Principles of traditional animation applied to 3d computer animation. *Proceedings of ACM SIGGRAPH*, 21(4):35–44.

Lazarus, R. S. (1991). *Emotion and Adaptation*. Oxford University Press.

Lazarus, R. S. (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual Review of Psychology*, 44:1–21.

Lee, S. P. (2002). *Facial Animation System with Realistic Eye Movement based on a Cognitive Model for Virtual Agents*. PhD thesis.

Lee, S. P., Badler, N. I., and Badler, J. (2002). Eyes alive. In *Proceedings of ACM SIGGRAPH*, pages 637–644.

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: The itc-sense of presence inventory. *Presence: Teleoperators and Virtual Environments*, 10(3):282–298.

Levenson, R. W., Ekman, P., and Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, 27(4):363–384.

Lim, M. Y., Aylett, R., and Jones, C. M. (2005). Affective guide with attitude. In Tao, J., Tan, T., and Picard, R. W., editors, *Proceedings of the First International Conference on Affective Computing and Intelligent Interaction*, volume 3784, pages 772–779. Springer-Verlag.

Lisetti, C. L. and Nasoz, F. (2002). Maui: a multimodal affective user interface. In *Multimedia*, pages 161–170.

Liu, C. K., Hertzmann, A., and Popović, Z. (2005). Learning physics-based motion style with nonlinear inverse optimization. *Proceedings of ACM SIGGRAPH*, 24(3):1071–1081.

Lombard, M. and Ditton, T. (1997). At the heart of it all: The concept of presence (available at http://www.ascusc.org/jcmc/vol3/issue2/lombard.html). *Journal of Computer-Mediated Communication*, 13(2).

Luborsky, L., Blinder, B., and Mackworth, N. (1963). Eye fixation and recall of pictures as a function of gsr responsivity. *Perceptual motor skills*, 16:469–483.

Machado, I., Paiva, A., and Prada, R. (2001). Is the wolf angry or... just hungry? In *Proceedings of the fifth international conference on Autonomous agents*, pages 370–376. ACM Press.

Marsella, S. and Gratch, J. (2001). Modeling the interplay of emotions and plans in multi-agent simulations. In *Cognitive Science Society*, volume Proceedings of the 23rd Annual Conference Cognitive Science Society.

Marsella, S. and Gratch, J. (2002). A step toward irrationality: using emotion to change belief. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 334–341. ACM Press.

Marsella, S., Lewis, J. W., and LaBore, C. (2000). Interactive pedagogical drama. In *Proceedings of Autonomous Agents*, volume 4th International Conference, pages 301–308.

MathWorks (2006). The mathworks - matlab and simulink for technical computing, http://www.mathworks.com/products/matlab/.

McCrae, R. R. and John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60:175–215.

McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models*. Chapman and Hall, second edition.

Meehan, M. (2001). *Physiological reaction as an objective measure of presence in virtual environments.* PhD thesis.

Meehan, M., Insko, B., Whitton, M. C., and Brooks, F. P. (2002). Physiological measures of presence in stressful virtual environments. In *Proceedings of ACM SIGGRAPH*, volume 21, pages 645–653.

Mehrabian, A. (1971). *Silent Messages.* Wadsworth Publishing Company, first edition.

Mehrabian, A. (1980). *Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies.* Oelgeschlager, Gunn & Hain.

Mehrabian, A. (1996a). Analysis of the big-five personality factors in terms of the pad temperament model. *Australian Journal of Psychology*, 48(2):86–92.

Mehrabian, A. (1996b). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14:261–292.

Mehrabian, A. and Firar, J. T. (1969). Encoding of attitude by a seated communicator via posture and position cues. *Journal of Consulting and Clinical Psychology*, 33(3):330–336.

Mignault, A. and Chaudhuri, A. (2003). The many faces of a neutral face: head tilt and perception of dominance and emotion. *Journal of Nonverbal Behavior*, 27(2):111–132.

Moffat, D. (1997). Personality parameters and programs. In *Creating Personalities for Synthetic Actors*, pages 120–165. Springer-Verlag.

Mogg, K. and Bradley, B. (1999). Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition and emotion*, 13(6):713–740.

Montepare, J., Koff, E., Zaitchik, D., and Albert, M. (1999). The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behaviour*, 23(2):133–152.

Mortensen, J., Vinayagamoorthy, V., Slater, M., Steed, A., Whitton, M. C., and Lok, B. (2002). Collaboration in tele-immersive environments. In *Proceedings of the Eight Eurographics Workshop on Virtual Environments*, pages 93–101.

Muirhead, R. D. and Goldman, M. (1979). Mutual eye contact as affected by seating position, sex and age. *Journal of Social Psychology*, 109:201–206.

Mukai, T. and Kuriyama, S. (2005). Geostatistical motion interpolation. *Proceedings of ACM SIGGRAPH*, 24(3):1062–1070.

Mulac, A., Studley, L. B., Wiemann, J. M., and Bradac, J. (1987). Male/female gaze in same-sex and mixed-sex dyads: gender-linked differences and mutual influence. *Human communication research*, 13:323–343.

Mullins, D. T. and Duke, M. P. (2004). Effects of social anxiety on nonverbal accuracy and response time i: facial expressions. *Journal of nonverbal behaviour*, 28(1):3–33.

NAG, T. N. A. G. (2006). Generalised linear interactive model (glim), http://www.nag.co.uk/stats/gdge_soft.asp/.

Nakano, Y. I., Reinstein, G., Stocky, T., and Cassell, J. (2003). Towards a model of face-to-face grounding. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics (ACL)*, pages 553–561. Association for Computational Linguistics.

Neumann, R., Seibt, B., and Strack, F. (2001). The influence of mood on the intensity of emotional responses: disentangling feeling and knowing. *Cognition and Emotion*, 15(6):725–747.

Nowak, K. L. and Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 12(5):481–494.

Oatley, K. and Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion*, 1:29–50.

Okino (2006). Nugraf home page - okino coputer graphics, http://www.okino.com/nrs/nrs.htm.

Ortony, A. (2003). *On making believable emotional agents believable*, chapter 6, pages 189–212. Emotions in Humans and Artifacts in Robert Trappl, Paolo Petta and Sabine Payr (eds.). Cambridge: MIT Press.

Ortony, A., Clore, G., and Collins, A. (1988). *Cognitive structure of emotions*. Cambridge University Press.

Ortony, A. and Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3):315–331.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobreperez, P., Woods, S., Zoll, C., and Hall, L. (2004). Caring for agents and agents that care: Building empathic relations with synthetic agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 194–201. IEEE Computer Society.

Paiva, A., Dias, J. a., Sobral, D., Aylett, R., Woods, S., Hall, L., and Zoll, C. (2005). Learn by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence*, 19(3–4):235–266.

Paterson, H. M., Pollick, F. E., and Sanford, A. J. (2001). The role of velocity in affect discrimination. In *J D Moore & K Stenning (Eds.) Twenty-Third annual conference of the cognitive science society*, pages 756–761. Lawrence ealrbaum associates, inc.

Patterson, M. L., Powell, J. L., and Lenihan, M. G. (1986). Touch, compliance and interpersonal affect. *Journal of Nonverbal Behaviour*, 10(1):41–50.

Pelachaud, C., Badler, N. I., and Steedman, M. (1994). Generating facial expressions for speech. Technical report.

Pelachaud, C. and Poggi, I. (2002). Subtleties of facial expressions in embodied agents. *Journal of Visualization and Computer Animation (Special issue on Graphical Autonomous Virtual Humans)*, 13(5):301–312.

Perlin, K. and Goldberg, A. (1996). Improv: A system for scripting interactive actors in virtual worlds. In *Proceedings of ACM SIGGRAPH*, pages 205–216.

Pertaub, D.-P., Slater, M., and Barker, C. (2001). An experiment on fear of public speaking in virtual reality. In *Medicine meets Virtual Reality*, pages 372–378.

Pertaub, D.-P., Slater, M., and Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments*, 11(1):68–78.

Peters, C. (2005). Direction of attention and interest using gaze behaviour. In Panayiotopoulos, T., Gratch, J., Aylett, R., Ballin, D., Olivier, P., and Rist, T., editors, *Proceedings of the 5th International Working Conference on Intelligent Virtual Agents*, volume 3661 of *Lecture Notes in Computer Science*, pages 215–228. Springer.

Peters, C., Pelachaud, C., Bevacqua, E., Mancini, M., and Poggi, I. (2005). A model of attention and interest using gaze beahviour. In Panayiotopoulos, T., Gratch, J., Aylett, R., Ballin, D., Olivier, P., and Rist, T., editors, *Proceedings of the 5th International Working Conference on Intelligent Virtual Agents*, volume 3661 of *Lecture Notes in Computer Science*, pages 229–239. Springer.

Pfungst, O. A. (1911). *Clever Hans (the Horse of Mr von Osten): a contribution to Experimental, Animal and Human Psychology (trans. C L Rahn).* Henry Holt.

Picard, R. W. (1997). *Affective computing.* MIT Press.

Planalp, S., DeFrancisco, V. L., and Rutherford, D. (1996). Varieties of cues to emotion in naturally occurring situations. *Cognition and Emotion*, 10(2):137–153.

Planalp, S. and Knie, K. (2002). *Integrating verbal and nonvrebal emotion(al) messages*, chapter 3, pages 55–77. The verbal communication of emotions: interdisciplinary perspectives. Lawrence Erlbaum Associates.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In Plutchik, R. and Kellerman, H., editors, *Emotion: Theory, research, and experience*, pages 3–33. Academic press, inc.

Poggi, I. and Pelachaud, C. (2002). *Signals and meanings of gaze in animated faces*, pages 133–144. Language, Vision and Music: Selected papers from the 8th International Workshop on the Cognitive Science of Natural Language Processing. John Benjamins.

Ponder, E. and Kennedy, W. P. (1927). On the act of blinking. *Quarterly journal of experimental psychology*, 18:89–110.

Prendinger, H. and Ishizuka, M. (2001). Let's talk! socially intelligent agents for language conversation training. In *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, volume 31, pages 465–471.

Prendinger, H., Mori, J., and Ishizuka, M. (2005). Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *International journal of human-computer studies*, 62(2):231–245.

Raouzaiou, A., Karpouzis, K., and Kollias, S. D. (2003). Online gaming and emotion representation. In *Proceedings of the International Workshop on Very Low Bitrate Video Coding (VLBV)*, pages 298–305.

Reeves, B. and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.* Cambridge university press.

Rehm, M. and André, E. (2005). Where do they look? gaze behaviors of multiple users interacting with an embodied conversational agent. In Panayiotopoulos, T., Gratch, J., Aylett, R., Ballin, D., Olivier, P., and Rist, T., editors, *Proceedings of the 5th International Workshop on Intelligent Virtual Agents (Lecture Notes in Computer Science)*, volume 3661, pages 241–252. Springer.

Reilly, S. N. (1997). A methodology for building believable social agents. In *Proceedings of the first international conference on Autonomous agents*, pages 114–121. ACM Press.

Rencher, A. C. (2002). *Methods of Multivariate Analysis.* Wiley-Interscience, second edition.

Richardson, D. C. and Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science: A Multidisciplinary Journal*, 29(6):1045–1060.

Rickel, J. and Lewis, J. W. (1998). Animated agents for procedural training in virtual reality: Perception, cognition and motor control. In *Applied Artificial Intelligence*, volume 13, pages 343–382.

Romano, D. M., Sheppard, G., Hall, J., Miller, A., and Ma, Z. (2005). Basic: A believable, adaptable socially intelligent character for social presence. In Slater, M., editor, *Proceedings of the 8th Annual International Workshop on Presence*, pages 81–96.

Rose, C., Cohen, M. F., and Bodenheimer, B. (1998). Verbs and adverbs: Multidimensional motion interpolation. *IEEE Computer Graphics and Applications*, 18(5):32–40.

Rosis, F. d., Pelachaud, C., Poggi, I., Carofiglio, V., and Carolis, B. D. (2003). From greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human-Computer Studies*, 59(1-2):81–118.

Russell, J. A. and Fernandez-Dols, J. M. (1997). *The psychology of facial expression.* Cambridge university press.

Sanchez-Vives, M. V. and Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, 6(4):332–339.

Schachter, S. (1959). *The psychology of affiliation*. Stanford University Press.

Schachter, S. and Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional states. *Psychological Review*, 69(5):379–399.

Scheflen, A. E. (1964). The significance of posture in communication systems. *Psychiatry*, 27:316–331.

Scheflen, A. E. and Scheflen, A. (1972). *Body language and social order: communication as behavioral control*. Prentice-Hall inc.

Scheirer, J., Fernandez, R., Klein, J., and Picard, R. W. (2002). Frustrating the user on purpose: a step toward building an affective computer. *Interacting with Computers*, 14(2):93–118.

Scheutz, M. (2004). Useful roles of emotions in artificial agents: A case study from artificial life. In *Proceedings of AAAI*. AAAI Press.

Scheutz, M., Sloman, A., and Logan, B. (2000). Emotional states and realistic agent behaviour. In *Proceedings GAME-ON*.

Schiano, D. J., Ehrlich, S. M., Rahardja, K., and Sheridan, K. (2000). Face to interface: facial affect in (hu)man and machine. In *Proceedings of ACM SIGCHI*, pages 193–200.

Schouwstra, S. J. and Hoogstraten, J. (1995). Head position and spinal position as determinants of perceived emotional state. *Perceptual and Motor Skills*, 81(2):673–674.

Schroeder, R. (1996). *Possible Worlds: The social dynamic of virtual reality technologies*. Westview Press.

Schroeder, R. (2002). *The Social Life of Avatars; Presence and Interaction in Shared Virtual Environments*. Springer-Verlag New York, Inc.

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281.

SICS (2006). The dive home page, swedish institute of computer science (sics), http://www.sics.se/dive.

Slater, M. (1999). Measuring presence: A response to the witmer and singer presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 8(5):560–566.

Slater, M. (2004). How colorful was your day?: Why questionnaires cannot assess presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(4):484–493.

Slater, M., Brogni, A., and Steed, A. (2003). Physiological responses to breaks in presence: A pilot study. In *Presence 2003*, volume The 6th Annual International Workshop on Presence.

Slater, M., Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., Brogni, A., Steed, A., and Friedman, D. (2006). Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 15(5).

Slater, M., Sadagic, A., Usoh, M., and Schroeder, R. (2000). Small group behaviour in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments*, 9(1):37–51.

Slater, M. and Steed, A. (2000). A virtual presence counter. *Presence: Teleoperators and Virtual Environments*, 9(2):214–217.

Slater, M. and Steed, A. (2002). *Meeting people virtually: experiments in shared virtual environments*, chapter 9, pages 146–171. The social life of avatars: presence and interaction in shared virtual environments. Springer-Verlag New York, Inc.

Slater, M., Steed, A., McCarthy, J., and Maringelli, F. (1998). The influence of body movement on subjective presence in virtual environments. *Human Factors*, 40(3):469–477.

Sloman, A. (2001a). Beyond shallow models of emotions. In *Cognitive Processing*, volume 2, pages 177–198.

Sloman, A. (2001b). Varieties of affect and the cogaff architecture schema. In *Proceedings of the Symposium on Emotion, Cognition, and Affective Computing - AISB Convention*, pages 39–48.

Smith, C. A. and Scott, H. S. (1997). A componential approach to the meaning of facial expressions. In Russel, J. A. and Fernández-Dols, J. M., editors, *The Psychology of Facial Expression*, pages 295–320. Cambridge University Press.

Sommer, R. (1969). *Personal space: The behavioural basis of design.* Prentice-Hall.

Sony (2006). Sony media software - sound forge 8 and sound forge audio studio, http://www.sonymediasoftware.com/products/soundforgefamily.asp.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., and Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory.* Consulting Psychologists Press, Palto, CA.

Steed, A., Brogni, A., and Vinayagamoorthy, V. (2005). Using breaks in presence to identify usability issues. In *The International Conference on Human Computer Interaction*, volume 11.

Steed, A., Mortensen, J., and Frécon, E. (2001). Spelunking: Experiences using the dive system on cave-like platform. *Immersive Projection Technologies and Virtual Environments*, pages 153–164.

Steed, A., Slater, M., Sadagic, A., and Tromp, J. G. (1999). Leadership and collaboration in virtual environments. In *IEEE Virtual Reality*, pages 58–63.

Stephenson, N. (1992). *Snow Crash.* ROC.

Tanguy, E. A. R., Bryson, J. J., and Willis, P. J. (2005). A dynamic emotion representation model within a facial animation system. Technical Report CSBU-2005-14, Dept. of Computer Science, Unversity of Bath.

Tashakkori, A. and Teddlie, C. (1998). *Mixed methodology: combining qualitative and quantitative approaches*, volume 46 of *Applied Social Research Methods Series*. Sage publications.

Taylor II, R. M., Hudson, T. C., Seeger, A., Weber, H., Juliano, J., and Helser, A. T. (2001). Vrpn: a device-independent, network-transparent vr peripheral system. In *Proceedings of the ACM symposium on Virtual Reality Software and Technology (VRST)*, pages 55–61, New York, NY, USA. ACM Press.

Thomas, F. and Johnson, O. (1981). *Disney Animation - The illusion of life*. Abbeville Press.

Thórisson, K. R. and Cassell, J. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. In *Applied Artificial Intelligence*, volume 13, pages 519–538.

Thorson, E. and Friestad, M. (1985). *The effects on emotion on episodic memory for television commercials*, pages 131–136. Advances in consumer psychology. Lexington.

Torres, O. E., Cassell, J., and Prevost, S. (1997). Modeling gaze behavior as a function of discourse structure. In *First International Workshop on Human Computer Conversations*.

Tromp, J. G., Bullock, A., Steed, A., Sadagic, A., Slater, M., and Frécon, E. (1998). Small group behaviour experiments in the coven project. In *IEEE Computer Graphics and Applications*, volume 18, pages 53–63.

Ushida, H., Hirayama, Y., and Nakajima, H. (1998). Emotion model for life-like agent and its evaluation. In *AAAI/IAAI: Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence*, pages 62–69. American Association for Artificial Intelligence.

Usoh, M., Catena, E., Arman, S., and Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments*, 9(5):497–503.

Velásquez, J. D. (1997). Modelling emotions and other motivations in synthetic agents. In *American Association for Artificial Intelligence*, pages 10–15.

Velásquez, J. D. (1998). A computational model of emotional conditioning in the brain. In *Grounding Emotions in Adaptive Systems at the 5th International Conference of the Society for Adaptive Behaviour (SAB)*.

Vilhjálmsson, H. H. (2005). Augmenting online conversation through automated discourse tagging. In *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS)*. IEEE Computer Society.

Vilhjálmsson, H. H. and Cassell, J. (1998). Bodychat: Autonomous communicative behaviours in avatars. In *International Conference on Autonomous Agents and multiagent systems*, volume 2, pages 269–276.

Vinayagamoorthy, V., Brogni, A., Gillies, M., Slater, M., and Steed, A. (2004a). An investigation of presence response across variations in visual realism. In *The 7th Annual International Presence Workshop*, pages 148–155.

Vinayagamoorthy, V., Brogni, A., Steed, A., and Slater, M. (2006a). The influence of a virtual character posture model on reported presence in an immersive virtual environment. In *ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications - VRCIA*, pages 229–236.

Vinayagamoorthy, V., Garau, M., Steed, A., and Slater, M. (2004b). An eye gaze model for dyadic interaction in an immersive virtual environment: Practice and experience. *Computer Graphics Forum*, 23(1):1–11.

Vinayagamoorthy, V., Gillies, M., Steed, A., Tanguy, E., Pan, X., Loscos, C., and Slater, M. (2006b). Building expression into virtual characters. In *Eurographics 2006 - State of the Art Report (STAR)*, pages 21–61.

Vinayagamoorthy, V., Steed, A., and Slater, M. (2005). Building characters: Lessons drawn from virtual environments. In *Towards Social Mechanisms of Android Science: A CogSci 2005 - XXVII Annual Conference of the Cognitive Science Society workshop*, pages 119–126.

Vyzas, E. and Picard, R. W. (1999). Offline and online recognition of emotion expression from physiological data. In *Workshop on Emotion-Based Agent Architectures at the Third International Conference on Autonomous Agents*, volume Technical Report No. 488.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, 28(6):879–896.

Walters, K. L. and Walk, R. D. (1986). Perception of emotion from body posture. *Bulletin of the Psychonomic Society*, 24(5):329–329.

Walters, K. L. and Walk, R. D. (1988). Perception of emotion from moving body cues in photographs. *Bulletin of the Psychonomic Society*, 26(2):112–114.

Watson, D. and Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, 33:448–457.

Watson, D. and Friend, R. (1987). *Social avoidance and distress scale [SAD] (1969) IN Corcoran K & Fischer J - Measures for clinical practice: A sourcebook.* Free Press, New york.

Watzlawick, P., Beavin, J. H., and Jackson, D. D. (1968). *Pragmatics of human communication.* Faber & Faber.

Wiener, M., Devoe, S., Robinson, S., and Geller, J. (1972). Nonverbal behaviour and nonverbal communication. *Psychological Review*, 79(3):185–214.

Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240.

Zhao, S. (2003). Toward a taxonomy of copresence. *Presence: Teleoperators and Virtual Environments*, 12(5):445–455.

Zimmons, P. and Panter, A. (2003). The influence of rendering quality on presence and task performance in a virtual environment. In *Proceedings of IEEE Virtual Reality*, pages 293–294.